National Institute for Health and Care Excellence

Draft for consultation

Osteoarthritis: assessment and management (update)

[I2] Evidence reviews for the clinical and cost-effectiveness of oral, topical and transdermal medicines for the management of osteoarthritis

NICE guideline

Evidence reviews underpinning recommendations 1.4.1 to 1.4.7 and research recommendations in the NICE guideline

April 2022

Draft for Consultation

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

1 Oral, topical and transdermal medicines for osteoarthritis

1.1.14 References

1. Abou-Raya S, Abou-Raya A, Helmii M. Duloxetine for the management of pain in older adults with knee osteoarthritis: randomised placebo-controlled trial. Age and Ageing. 2012; 41(5):646-652

2. Afilalo M, Etropolski MS, Kuperwasser B, Kelly K, Okamoto A, Van Hove I et al. Efficacy and safety of Tapentadol extended release compared with oxycodone controlled release for the management of moderate to severe chronic pain related to osteoarthritis of the knee: a randomized, double-blind, placebo- and active-controlled phase III study. Clinical Drug Investigation. 2010; 30(8):489-505

3. Altman RD, Dreiser RL, Fisher CL, Chase WF, Dreher DS, Zacher J. Diclofenac sodium gel in patients with primary hand osteoarthritis: a randomized, double-blind, placebo-controlled trial. Journal of Rheumatology. 2009; 36(9):1991-1999

4. Altman RD, Zinsenheim JR, Temple AR, Schweinle JE. Three-month efficacy and safety of acetaminophen extended-release for osteoarthritis pain of the hip or knee: a randomized, double-blind, placebo-controlled study. Osteoarthritis and Cartilage. 2007; 15(4):454-461

5. Ammendolia A, Marotta N, Marinaro C, Demeco A, Mondardini P, Costantino C. The synergic use of the high power laser therapy and glu-cosamine sulfate in knee osteoarthritis: A randomized controlled trial. Acta Biomedica. 2021; 92 (3) (no pagination)(e2021237)

6. Amundsen T, Bleken L, Borkje B. Variation in response to naproxen and diclofenac in patients with osteoarthritis. Current therapeutic research - clinical and experimental. 1983; 33(5):793-801

7. Andelman SY. Etodolac, aspirin, and placebo in patients with degenerative joint disease: a twelve-week study. Clinical Therapeutics. 1983; 5(6):651-661

8. Anonymous. Treatment of osteoarthritis with piroxicam. European Journal of Rheumatology and Inflammation. 1983; 6(1):73-83

9. Asmus MJ, Essex MN, Brown PB, Mallen SR. Efficacy and tolerability of celecoxib in osteoarthritis patients who previously failed naproxen and ibuprofen: results from two trials. International Journal of Clinical Rheumatology. 2014; 9(6):551-558

10. Asmus MJ, Essex MN, Brownz PB, Mallen SR. Efficacy and tolerability of celecoxib in patients with osteoarthritis who previously did not respond to or did not tolerate naproxen and ibuprofen: results from 2 identically designed randomized trials. Osteoarthritis and cartilage. 2013; 21:S253

11. Babul N, Noveck R, Chipman H, Roth SH, Gana T, Albert K. Efficacy and safety of extended-release, once-daily tramadol in chronic pain: a randomized 12-week clinical trial in osteoarthritis of the knee. Journal of Pain and Symptom Management. 2004; 28(1):59-71

12. Baer PA, Thomas LM, Shainhouse Z. Treatment of osteoarthritis of the knee with a topical diclofenac solution: a randomised controlled, 6-week trial [ISRCTN53366886]. BMC Musculoskeletal Disorders. 2005; 6:44

13. Baerwald C, Verdecchia P, Duquesroix B, Frayssinet H, Ferreira T. Efficacy, safety, and effects on blood pressure of naproxcinod 750 mg twice daily compared with placebo and naproxen 500 mg twice daily in patients with osteoarthritis of the hip: a randomized, double-blind, parallel-group, multicenter study. Arthritis and Rheumatism. 2010; 62(12):3635-3644

14. Bakshi R, Darekar B, Langdon CG, Rotman H. Efficacy and tolerability of diclofenac dispersible in elderly patients with osteoarthritis. Current Medical Research and Opinion. 1991; 12(7):459-465

15. Banerjee M, Mondal S, Sarkar R, Mondal H, Bhattacharya K. Comparative study of efficacy and safety of tapentadol versus etoricoxib in mild to moderate grades of chronic osteorthritis of knee. 2016; 11(1):21-25

16. Baraf HS, Gold MS, Clark MB, Altman RD. Safety and efficacy of topical diclofenac sodium 1% gel in knee osteoarthritis: a randomized controlled trial. Physician & Sportsmedicine. 2010; 38(2):19-28

17. Barthel HR, Haselwood D, Longley S, 3rd, Gold MS, Altman RD. Randomized controlled trial of diclofenac sodium gel in knee osteoarthritis. Seminars in Arthritis and Rheumatism. 2009; 39(3):203-212

18. Batlle-Gualda E, Roman Ivorra J, Martin-Mola E, Carbonell Abello J, Linares Ferrando LF, Tornero Molina J et al. Aceclofenac vs paracetamol in the management of symptomatic osteoarthritis of the knee: a double-blind 6-week randomized controlled trial. Osteoarthritis and Cartilage. 2007; 15(8):900-908

19. Beaulieu AD, Peloso PM, Haraoui B, Bensen W, Thomson G, Wade J et al. Once-daily, controlled-release tramadol and sustained-release diclofenac relieve chronic pain due to osteoarthritis: a randomized controlled trial. Pain Research & Management. 2008; 13(2):103-110

20. Bensen WG, Fiechtner JJ, McMillen JI, Zhao WW, Yu SS, Woods EM et al. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. Mayo Clinic Proceedings. 1999; 74(11):1095-1105

21. Berry H, Bloom B, Hamilton EBD. A comparative study of piroxicam (Feldene), diclofenac and placebo in osteoarthritis. Clinical Trials Journal. 1982; 19(6):349-358

22. Berry H, Coquelin JP, Gordon A, Seymour D. Antrafenine, naproxen and placebo in osteoarthritis: a comparative study. British Journal of Rheumatology. 1983; 22(2):89-94

23. Bhatia A, Goni V, Chopra S, Singh B, Katare OP. Evaluation of efficacy and safety of a novel lipogel containing diclofenac: A randomized, placebo controlled, double-blind clinical trial in patients with signs and symptoms of osteoarthritis. Contemporary Clinical Trials Communications. 2020; 20:100664

24. Bingham CO, 3rd, Bird SR, Smugar SS, Xu X, Tershakovec AM. Responder analysis and correlation of outcome measures: pooled results from two identical studies comparing etoricoxib, celecoxib, and placebo in osteoarthritis. Osteoarthritis and Cartilage. 2008; 16(11):1289-1293

25. Bingham CO, 3rd, Sebba AI, Rubin BR, Ruoff GE, Kremer J, Bird S et al. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies. Rheumatology. 2007; 46(3):496-507

26. Bingham CO, 3rd, Smugar SS, Wang H, Peloso PM, Gammaitoni A. Predictors of response to cyclo-oxygenase-2 inhibitors in osteoarthritis: pooled results from two

identical trials comparing etoricoxib, celecoxib, and placebo. Pain Medicine. 2011; 12(3):352-361

27. Bingham CO, 3rd, Smugar SS, Wang H, Tershakovec AM. Early response to COX-2 inhibitors as a predictor of overall response in osteoarthritis: pooled results from two identical trials comparing etoricoxib, celecoxib and placebo. Rheumatology. 2009; 48(9):1122-1127

28. Biondi DM, Xiang J, Etropolski M, Moskovitz B. Tolerability and efficacy of tapentadol extended release in elderly patients >= 75 years of age with chronic osteoarthritis knee or low back pain. Journal of Opioid Management. 2015; 11(5):393-403

29. Birbara C, Ruoff G, Sheldon E, Valenzuela C, Rodgers A, Petruschke RA et al. Efficacy and safety of rofecoxib 12.5 mg and celecoxib 200 mg in two similarly designed osteoarthritis studies. Current Medical Research and Opinion. 2006; 22(1):199-210

30. Blondi D, Xiang J, Lange R, Etropolski M, Vorsanger G, Moskovitz B. Tolerability of tapentadol extended release versus oxycodone controlled release in elderly patients with chronic low back or osteoarthritis pain in a 1-year safety study. Der schmerz. 2010; 24(S1):92

31. Bocanegra TS, Weaver AL, Tindall EA, Sikes DH, Ball JA, Wallemark CB et al. Diclofenac/misoprostol compared with diclofenac in the treatment of osteoarthritis of the knee or hip: a randomized, placebo controlled trial. Arthrotec Osteoarthritis Study Group. Journal of Rheumatology. 1998; 25(8):1602-1611

32. Bolten W, Gomes JA, Stead H, Geis GS. The gastroduodenal safety and efficacy of the fixed combination of diclofenac and misoprostol in the treatment of osteoarthritis. British Journal of Rheumatology. 1992; 31(11):753-758

33. Bookman AA, Williams KS, Shainhouse JZ. Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. CMAJ Canadian Medical Association Journal. 2004; 171(4):333-338

34. Boureau F, Schneid H, Zeghari N, Wall R, Bourgeois P. The IPSO study: ibuprofen, paracetamol study in osteoarthritis. A randomised comparative clinical study comparing the efficacy and safety of ibuprofen and paracetamol analgesic treatment of osteoarthritis of the knee or hip. Annals of the Rheumatic Diseases. 2004; 63(9):1028-1034

35. Bradley JD, Brandt KD, Katz BP, Kalasinski LA, Ryan SI. Comparison of an antiinflammatory dose of ibuprofen, an analgesic dose of ibuprofen, and acetaminophen in the treatment of patients with osteoarthritis of the knee. New England Journal of Medicine. 1991; 325(2):87-91

36. Bradley JD, Katz BP, Brandt KD. Severity of knee pain does not predict a better response to an antiinflammatory dose of ibuprofen than to analgesic therapy in patients with osteoarthritis. Journal of Rheumatology. 2001; 28(5):1073-1076

37. Breivik H, Ljosaa TM, Stengaard-Pedersen K, Persson J, Aro H, Villumsen J et al. A 6-months, randomised, placebo-controlled evaluation of efficacy and tolerability of a low-dose 7-day buprenorphine transdermal patch in osteoarthritis patients naive to potent opioids. Scandinavian Journal of Pain. 2010; 1(3):122-141

38. Bruyere O, Pavelka K, Rovati LC, Deroisy R, Olejarova M, Gatterova J et al. Glucosamine sulfate reduces osteoarthritis progression in postmenopausal women with knee osteoarthritis: evidence from two 3-year studies. Menopause. 2004; 11(2):138-143

39. Burch F, Fishman R, Messina N, Corser B, Radulescu F, Sarbu A et al. A comparison of the analgesic efficacy of Tramadol Contramid OAD versus placebo in patients with pain due to osteoarthritis. Journal of Pain and Symptom Management. 2007; 34(3):328-338

40. Cahlin BJ, Dahlstrom L. No effect of glucosamine sulfate on osteoarthritis in the temporomandibular joints--a randomized, controlled, short-term study. Oral Surgery Oral Medicine Oral Pathology Oral Radiology & Endodontics. 2011; 112(6):760-766

41. Caldwell JR, Rapoport RJ, Davis JC, Offenberg HL, Marker HW, Roth SH et al. Efficacy and safety of a once-daily morphine formulation in chronic, moderate-to-severe osteoarthritis pain: results from a randomized, placebo-controlled, double-blind trial and an open-label extension trial. Journal of Pain and Symptom Management. 2002; 23(4):278-291

42. Caroit M, Forette B, Hubault A, Pasquier P. Double blind study of ketoprofen against a placebo in osteoarthritis of the hip. Scandinavian Journal of Rheumatology. 1976; 5(sup.14):123-127

43. Case JP, Baliunas AJ, Block JA. Lack of efficacy of acetaminophen in treating symptomatic knee osteoarthritis: a randomized, double-blind, placebo-controlled comparison trial with diclofenac sodium. Archives of Internal Medicine. 2003; 163(2):169-178

44. Chan FK, Lanas A, Scheiman J, Berger MF, Nguyen H, Goldstein JL. Celecoxib versus omeprazole and diclofenac in patients with osteoarthritis and rheumatoid arthritis (CONDOR): a randomised trial. Lancet. 2010; 376(9736):173-179

45. Chappell AS, Desaiah D, Liu-Seifert H, Zhang S, Skljarevski V, Belenkov Y et al. A double-blind, randomized, placebo-controlled study of the efficacy and safety of duloxetine for the treatment of chronic pain due to osteoarthritis of the knee. Pain Practice. 2011; 11(1):33-41

46. Chappell AS, Ossanna MJ, Liu-Seifert H, Iyengar S, Skljarevski V, Li LC et al. Duloxetine, a centrally acting analgesic, in the treatment of patients with osteoarthritis knee pain: a 13-week, randomized, placebo-controlled trial. Pain. 2009; 146(3):253-260

47. Chindalore VL, Craven RA, Yu KP, Butera PG, Burns LH, Friedmann N. Adding ultralow-dose naltrexone to oxycodone enhances and prolongs analgesia: a randomized, controlled trial of Oxytrex. Journal of Pain. 2005; 6(6):392-399

48. Chopra A, Saluja M, Tillu G, Sarmukkaddam S, Venugopalan A, Narsimulu G et al. Ayurvedic medicine offers a good alternative to glucosamine and celecoxib in the treatment of symptomatic knee osteoarthritis: a randomized, double-blind, controlled equivalence drug trial. Rheumatology. 2013; 52(8):1408-1417

49. Cibere J, Kopec JA, Thorne A, Singer J, Canvin J, Robinson DB et al. Randomized, double-blind, placebo-controlled glucosamine discontinuation trial in knee osteoarthritis. Arthritis and Rheumatism. 2004; 51(5):738-745

50. Clegg DO, Reda DJ, Harris CL, Klein MA, O'Dell JR, Hooper MM et al. Glucosamine, chondroitin sulfate, and the two in combination for painful knee osteoarthritis. New England Journal of Medicine. 2006; 354(8):795-808

51. Conaghan PG, Dickson J, Bolten W, Cevc G, Rother M. A multicentre, randomized, placebo- and active-controlled trial comparing the efficacy and safety of topical ketoprofen in Transfersome gel (IDEA-033) with ketoprofen-free vehicle (TDT 064)

and oral celecoxib for knee pain associated with osteoarthritis. Rheumatology. 2013; 52(7):1303-1312

52. Couto A, Troullos E, Moon J, Paredes-Diaz A, An R. Analgesic efficacy and safety of non-prescription doses of naproxen sodium in the management of moderate osteoarthritis of the knee or hip. Current Medical Research and Opinion. 2018; 34(10):1747-1753

53. Cryer BL, Sostek MB, Fort JG, Svensson O, Hwang C, Hochberg MC. A fixed-dose combination of naproxen and esomeprazole magnesium has comparable upper gastrointestinal tolerability to celecoxib in patients with osteoarthritis of the knee: results from two randomized, parallel-group, placebo-controlled trials. Annals of Medicine. 2011; 43(8):594-605

54. Dehghan M, Saffari M, Rafieian-kopaei M, Ahmadi A, Lorigooini Z. Comparison of the effect of topical Hedera helix L. extract gel to diclofenac gel in the treatment of knee osteoarthritis. Journal of Herbal Medicine. 2020; 22 (no pagination)(100350)

55. DeLemos BP, Xiang J, Benson C, Gana TJ, Pascual ML, Rosanna R et al. Tramadol hydrochloride extended-release once-daily in the treatment of osteoarthritis of the knee and/or hip: a double-blind, randomized, dose-ranging trial. American Journal of Therapeutics. 2011; 18(3):216-226

56. Dickson DJ. A double-blind evaluation of topical piroxicam gel with oral ibuprofen in osteoarthritis of the knee. Current therapeutic research. 1991; 49:199-107

57. Dieppe P, Cushnaghan J, Jasani MK, McCrae F, Watt I. A two-year, placebo-controlled trial of non-steroidal anti-inflammatory therapy in osteoarthritis of the knee joint. British Journal of Rheumatology. 1993; 32(7):595-600

58. Doherty M, Hawkey C, Goulder M, Gibb I, Hill N, Aspley S et al. A randomised controlled trial of ibuprofen, paracetamol or a combination tablet of ibuprofen/paracetamol in community-derived people with knee pain. Annals of the Rheumatic Diseases. 2011; 70(9):1534-1541

59. Dore R, Ballard I, Constantine G, McDonald P. Efficacy and safety of etodolac and naproxen in patients with osteoarthritis of the knee: a double-blind, placebo-controlled study. Clinical Therapeutics. 1995; 17(4):656-666

60. Enteshari-Moghaddam A, Azami A, Isazadehfar K, Mohebbi H, Habibzadeh A, Jahanpanah P. Efficacy of duloxetine and gabapentin in pain reduction in patients with knee osteoarthritis. Clinical Rheumatology. 2019; 38(10):2873-2880

61. Esselinckx W, Stenier P. Double-blind comparative study of tenoxicam in arthrosis of the knee and hip. Current Therapeutic Research, Clinical and Experimental. 1990; 48(1):206-215

62. Essex MN, Behar R, O'Connell MA, Brown PB. Efficacy and tolerability of celecoxib and naproxen versus placebo in Hispanic patients with knee osteoarthritis. International Journal of General Medicine. 2014; 7:227-235

63. Essex MN, O'Connell M, Bhadra Brown P. Response to nonsteroidal anti-inflammatory drugs in African Americans with osteoarthritis of the knee. Journal of International Medical Research. 2012; 40(6):2251-2266

64. Essex MN, O'Connell MA, Behar R, Bao W. Efficacy and safety of nonsteroidal anti-inflammatory drugs in Asian patients with knee osteoarthritis: summary of a randomized, placebo-controlled study. International Journal of Rheumatic Diseases. 2016; 19(3):262-270

65. Etropolski M, Kuperwasser B, Flugel M, Haufel T, Lange B, Rauschkolb C et al. Safety and tolerability of tapentadol extended release in moderate to severe chronic osteoarthritis or low back pain management: pooled analysis of randomized controlled trials. Advances in Therapy. 2014; 31(6):604-620

66. Famaey JP, Colinet E. A double blind trial of ketoprofen in the treatment of osteoarthritis of the hip. Rheumatology. 1976; 15(Suppl):45-49

67. Fishman RL, Kistler CJ, Ellerbusch MT, Aparicio RT, Swami SS, Shirley ME et al. Efficacy and safety of 12 weeks of osteoarthritic pain therapy with once-daily tramadol (Tramadol Contramid OAD). Journal of Opioid Management. 2007; 3(5):273-280

68. Fleischmann R, Sheldon E, Maldonado-Cocco J, Dutta D, Yu S, Sloan VS. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a prospective randomized 13-week study versus placebo and celecoxib. Clinical Rheumatology. 2006; 25(1):42-53

69. Fleischmann RM, Caldwell JR, Roth SH, Tesser JRP, Olson W, Kamin M. Tramadol for the treatment of joint pain associated with osteoarthritis: a randomized, double-blind, placebo-controlled trial. Current therapeutic research - clinical and experimental. 2001; 62(2):113-128

70. Fleischmann RM, Flint K, Constantine G, Kolecki B. A double-masked comparison of Naprelan and nabumetone in osteoarthritis of the knee. Naprelan Study Group. Clinical Therapeutics. 1997; 19(4):642-655

71. Florete OG, Xiang J, Vorsanger GJ. Effects of extended-release tramadol on pain-related sleep parameters in patients with osteoarthritis. Expert Opinion on Pharmacotherapy. 2008; 9(11):1817-1827

72. Frakes EP, Risser RC, Ball TD, Hochberg MC, Wohlreich MM. Duloxetine added to oral nonsteroidal anti-inflammatory drugs for treatment of knee pain due to osteoarthritis: results of a randomized, double-blind, placebo-controlled trial. Current Medical Research and Opinion. 2011; 27(12):2361-2372

73. Fransen M, Agaliotis M, Nairn L, Votrubec M, Bridgett L, Su S et al. Glucosamine and chondroitin for knee osteoarthritis: a double-blind randomised placebo-controlled clinical trial evaluating single and combination regimens. Annals of the Rheumatic Diseases. 2015; 74(5):851-858

74. Frestedt JL, Walsh M, Kuskowski MA, Zenk JL. A natural mineral supplement provides relief from knee osteoarthritis symptoms: a randomized controlled pilot trial. Nutrition Journal. 2008; 7:9

75. Frestedt JL, Walsh M, Kuskowski MA, Zenk JL. The natural mineral supplement, Aquamin, provides relief from the symptoms of knee osteoarthritis symptoms: A randomised controlled pilot trial. Osteoarthritis and Cartilage. 2009; 1):S174

76. Friedmann N, Klutzaritz V, Webster L. Efficacy and safety of an extended-release oxycodone (Remoxy) formulation in patients with moderate to severe osteoarthritic pain. Journal of Opioid Management. 2011; 7(3):193-202

77. Gana TJ, Pascual ML, Fleming RR, Schein JR, Janagap CC, Xiang J et al. Extended-release tramadol in the treatment of osteoarthritis: a multicenter, randomized, double-blind, placebo-controlled clinical trial. Current Medical Research and Opinion. 2006; 22(7):1391-1401

78. Geba GP, Weaver AL, Polis AB, Dixon ME, Schnitzer TJ, Vioxx ACTG. Efficacy of rofecoxib, celecoxib, and acetaminophen in osteoarthritis of the knee: a randomized trial. JAMA. 2002; 287(1):64-71

79. Ghosh S, Paul S, Das N, Bhattacharyya TK. A study on the effects of diclofenac sodium and etoricoxib in the treatment of osteoarthritis. Journal of the Indian Medical Association. 2007; 105(5):260-262

80. Giansiracusa JE, Donaldson MS, Koonce ML, Lefton TE, Ruoff GE, Brooks CD. Ibuprofen in osteoarthritis. Southern Medical Journal. 1977; 70(1):49-52

81. Gibofsky A, Hochberg MC, Jaros MJ, Young CL. Efficacy and safety of low-dose submicron diclofenac for the treatment of osteoarthritis pain: a 12 week, phase 3 study. Current Medical Research and Opinion. 2014; 30(9):1883-1893

82. Gibofsky A, Williams GW, McKenna F, Fort JG. Comparing the efficacy of cyclooxygenase 2-specific inhibitors in treating osteoarthritis: appropriate trial design considerations and results of a randomized, placebo-controlled trial. Arthritis and Rheumatism. 2003; 48(11):3102-3111

83. Giordano N, Fioravanti A, Papakostas P, Montella A, Giorgi G, Nuti R. The efficacy and tolerability of glucosamine sulfate in the treatment of knee osteoarthritis: A randomized, double-blind, placebo-controlled trial. Current Therapeutic Research, Clinical and Experimental. 2009; 70(3):185-196

84. Golden HE, Moskowitz RW, Minic M. Analgesic efficacy and safety of nonprescription doses of naproxen sodium compared with acetaminophen in the treatment of osteoarthritis of the knee. American Journal of Therapeutics. 2004; 11(2):85-94

85. Goldstein JL, Cryer B, Amer F, Hunt B. Celecoxib plus aspirin versus naproxen and lansoprazole plus aspirin: a randomized, double-blind, endoscopic trial. Clinical Gastroenterology and Hepatology. 2007; 5(10):1167-1174

86. Gordo AC, Walker C, Armada B, Zhou D. Efficacy of celecoxib versus ibuprofen for the treatment of patients with osteoarthritis of the knee: A randomized double-blind, non-inferiority trial. Journal of International Medical Research. 2017; 45(1):59-74

87. Gottesdiener K, Schnitzer T, Fisher C, Bockow B, Markenson J, Ko A et al. Results of a randomized, dose-ranging trial of etoricoxib in patients with osteoarthritis. Rheumatology. 2002; 41(9):1052-1061

88. Grace D, Rogers J, Skeith K, Anderson K. Topical diclofenac versus placebo: a double blind, randomized clinical trial in patients with osteoarthritis of the knee. Journal of Rheumatology. 1999; 26(12):2659-2663

89. Haghighi M, Khalvat A, Toliat T, Jallaei S. Comparing the effects of ginger (Zingiber officinale) extract and ibuprofen on patients with osteoarthritis. Archives of Iranian Medicine. 2005; 8(4):267-271

90. Herrero-Beaumont G, Ivorra JA, Del Carmen Trabado M, Blanco FJ, Benito P, Martin-Mola E et al. Glucosamine sulfate in the treatment of knee osteoarthritis symptoms: a randomized, double-blind, placebo-controlled study using acetaminophen as a side comparator. Arthritis and Rheumatism. 2007; 56(2):555-567

91. Hochberg MC, Clegg DO. Potential effects of chondroitin sulfate on joint swelling: a GAIT report. Osteoarthritis and Cartilage. 2008; 16 Suppl 3:S22-24

92. Hochberg MC, Fort JG, Svensson O, Hwang C, Sostek M. Fixed-dose combination of enteric-coated naproxen and immediate-release esomeprazole has comparable

efficacy to celecoxib for knee osteoarthritis: two randomized trials. Current Medical Research and Opinion. 2011; 27(6):1243-1253

93. Houpt JB, McMillan R, Wein C, Paget-Dellio SD. Effect of glucosamine hydrochloride in the treatment of pain of osteoarthritis of the knee. Journal of Rheumatology. 1999; 26(11):2423-2430

94. Hubault A, Caroit M, Forette B, Pasquier P. Double blind trial of ketoprofen compared with placebo in osteoarthrosis of the hip. Rheumatology. 1976; 15(Suppl):52-55

95. Hudson B, Williman JA, Stamp LK, Alchin JS, Hooper GJ, Mangin D et al. Nortriptyline for pain in knee osteoarthritis: a double-blind randomised controlled trial in New Zealand general practice. British Journal of General Practice. 2021; 71(708):e538-e546

96. Hughes R, Carr A. A randomized, double-blind, placebo-controlled trial of glucosamine sulphate as an analgesic in osteoarthritis of the knee. Rheumatology. 2002; 41(3):279-284

97. Kageyama T. Clinical evaluation of naproxen in the treatment of osteoarthritis--double-blind, cross-over trial. Scandinavian Journal of Rheumatology - Supplement. 1973; 2:94-100

98. Karakaya M. A double-blind comparison of a new anti-inflammatory stubstance, proquazone, with indomethacin and placebo in osteoarthritis of the knee joint. Current therapeutic research clinical and experimental. 1977; 22(1):127

99. Karlsson M, Berggren AC. Efficacy and safety of low-dose transdermal buprenorphine patches (5, 10, and 20 microg/h) versus prolonged-release tramadol tablets (75, 100, 150, and 200 mg) in patients with chronic osteoarthritis pain: a 12-week, randomized, open-label, controlled, parallel-group noninferiority study. Clinical Therapeutics. 2009; 31(3):503-513

100. Kean WF, Bouchard S, Roderich Gossen E. Women with pain due to osteoarthritis: the efficacy and safety of a once-daily formulation of tramadol. Pain Medicine. 2009; 10(6):1001-1011

101. Kellner HL, Li C, Essex MN. Efficacy and safety of celecoxib versus diclofenac and omeprazole in elderly arthritis patients: a subgroup analysis of the CONDOR trial. Current Medical Research and Opinion. 2012; 28(9):1537-1545

102. Kivitz A, Eisen G, Zhao WW, Bevirt T, Recker DP. Randomized placebo-controlled trial comparing efficacy and safety of valdecoxib with naproxen in patients with osteoarthritis. Journal of Family Practice. 2002; 51(6):530-537

103. Kivitz AJ, Greenwald MW, Cohen SB, Polis AB, Najarian DK, Dixon ME et al. Efficacy and safety of rofecoxib 12.5 mg versus nabumetone 1,000 mg in patients with osteoarthritis of the knee: a randomized controlled trial. Journal of the American Geriatrics Society. 2004; 52(5):666-674

104. Kivitz AJ, Makarowski WS, Fiechtner JJ, Recker DP. A flexible daily dosage regimen of oxaprozin potassium in patients with acute knee pain associated with osteoarthritis: 24-Hour analgesic durability and safety. Clinical Drug Investigation. 2001; 21(11):745-753

105. Kivitz AJ, Moskowitz RW, Woods E, Hubbard RC, Verburg KM, Lefkowith JB et al. Comparative efficacy and safety of celecoxib and naproxen in the treatment of osteoarthritis of the hip. Journal of International Medical Research. 2001; 29(6):467-479

106. Kneer W, Rother M, Mazgareanu S, Seidel EJ, European I-sg. A 12-week randomized study of topical therapy with three dosages of ketoprofen in Transfersome gel (IDEA-033) compared with the ketoprofen-free vehicle (TDT 064), in patients with osteoarthritis of the knee. Journal of Pain Research. 2013; 6:743-753

107. Kosinski M, Janagap C, Gajria K, Schein J, Freedman J. Pain relief and pain-related sleep disturbance with extended-release tramadol in patients with osteoarthritis. Current Medical Research and Opinion. 2007; 23(7):1615-1626

108. Kosuwon W, Sirichatiwapee W, Wisanuyotin T, Jeeravipoolvarn P, Laupattarakasem W. Efficacy of symptomatic control of knee osteoarthritis with 0.0125% of capsaicin versus placebo. Journal of the Medical Association of Thailand. 2010; 93(10):1188-1195

109. Kwoh CK, Roemer FW, Hannon MJ, Moore CE, Jakicic JM, Guermazi A et al. Effect of oral glucosamine on joint structure in individuals with chronic knee pain: a randomized, placebo-controlled clinical trial. Arthritis & Rheumatology. 2014; 66(4):930-939

110. Laine L, Harper S, Simon T, Bath R, Johanson J, Schwartz H et al. A randomized trial comparing the effect of rofecoxib, a cyclooxygenase 2-specific inhibitor, with that of ibuprofen on the gastroduodenal mucosa of patients with osteoarthritis. Rofecoxib Osteoarthritis Endoscopy Study Group. Gastroenterology. 1999; 117(4):776-783

111. Lange B, Kuperwasser B, Okamoto A, Steup A, Haufel T, Ashworth J et al. Efficacy and safety of tapentadol prolonged release for chronic osteoarthritis pain and low back pain. Advances in Therapy. 2010; 27(6):381-399

112. Lange B, Sohns M, Tempero J, Elling C. Efficacy and safety of tapentadol prolonged release formulation in the treatment of elderly patients with moderate-to-severe chronic osteoarthritis knee pain: a pooled analysis of two double-blind, randomized, placebo-, and active-controlled trials. Current Medical Research and Opinion. 2018; 34(12):2113-2123

113. Lange B, von Zabern D, Elling C, Dubois C. Efficacy and safety of tapentadol prolonged release for moderate-to-severe chronic osteoarthritis knee pain: a pooled analysis of two double-blind, randomized, placebo- and oxycodone controlled release-controlled studies. Current Medical Research and Opinion. 2017; 33(8):1413-1422

114. Langford R, McKenna F, Ratcliffe S, Vojtassak J, Richarz U. Transdermal fentanyl for improvement of pain and functioning in osteoarthritis: a randomized, placebo-controlled trial. Arthritis and Rheumatism. 2006; 54(6):1829-1837

115. Leatham PA, Bird HA, Wright V, Seymour D, Gordon A. A double blind study of antrafenine, naproxen and placebo in osteoarthrosis. European Journal of Rheumatology and Inflammation. 1983; 6(2):209-211

116. Lee M, Yoo J, Kim JG, Kyung HS, Bin SI, Kang SB et al. A Randomized, Multicenter, Phase III Trial to Evaluate the Efficacy and Safety of Polmacoxib Compared with Celecoxib and Placebo for Patients with Osteoarthritis. Clinics in Orthopedic Surgery. 2017; 9(4):439-457

117. Lehmann R, Brzosko M, Kopsa P, Nischik R, Kreisse A, Thurston H et al. Efficacy and tolerability of lumiracoxib 100 mg once daily in knee osteoarthritis: a 13-week, randomized, double-blind study vs. placebo and celecoxib. Current Medical Research and Opinion. 2005; 21(4):517-526

118. Leigh TJ, Bird HA, Hindmarch I, Wright V. The evaluation of a nonsteroidal anti-inflammatory drug (tenoxicam) in the treatment of sleep disturbance in osteoarthritic patients. Clinical Rheumatology. 1989; 8(4):489-493

119. Leung AT, Malmstrom K, Gallacher AE, Sarembock B, Poor G, Beaulieu A et al. Efficacy and tolerability profile of etoricoxib in patients with osteoarthritis: A randomized, double-blind, placebo and active-comparator controlled 12-week efficacy trial. Current Medical Research and Opinion. 2002; 18(2):49-58

120. Lohmander LS, McKeith D, Svensson O, Malmenas M, Bolin L, Kalla A et al. A randomised, placebo controlled, comparative trial of the gastrointestinal safety and efficacy of AZD3582 versus naproxen in osteoarthritis. Annals of the Rheumatic Diseases. 2005; 64(3):449-456

121. Lopes Vaz A. Double-blind clinical evaluation of the relative efficacy of ibuprofen and glucosamine sulphate in the management of osteoarthrosis of the knee in out-patients. Current Medical Research and Opinion. 1982; 8(3):145-149

122. Lopez Sanchez J, Yanez Marchena G. Meclofenamate sodium, phenylbutazone, and naproxen in the treatment of degenerative joint disease. Report of a placebo-controlled double-blind clinical comparison. Arzneimittel-Forschung/Drug Research. 1983; 33(4 A):653-656

123. Lund B, Distel M, Bluhmki E. A double-blind, randomized, placebo-controlled study of efficacy and tolerance of meloxicam treatment in patients with osteoarthritis of the knee. Scandinavian Journal of Rheumatology. 1998; 27(1):32-37

124. Makarowski W, Weaver A, Rubin B, Caldwell J, McMahon FG, Noveck RJ et al. The efficacy, tolerability, and safety of 1200 mg/d of oxaprozin and 1500 mg/d of nabumetone in the treatment of patients with osteoarthritis of the knee. Clinical Therapeutics. 1996; 18(1):114-124

125. Makarowski W, Zhao WW, Bevirt T, Recker DP. Efficacy and safety of the COX-2 specific inhibitor valdecoxib in the management of osteoarthritis of the hip: a randomized, double-blind, placebo-controlled comparison with naproxen. Osteoarthritis and Cartilage. 2002; 10(4):290-296

126. Malonne H, Coffiner M, Sonet B, Sereno A, Vanderbist F. Efficacy and tolerability of sustained-release tramadol in the treatment of symptomatic osteoarthritis of the hip or knee: a multicenter, randomized, double-blind, placebo-controlled study. Clinical Therapeutics. 2004; 26(11):1774-1782

127. March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. BMJ. 1994; 309(6961):1041-1045; discussion 1045-1046

128. March L, Schwarz J, Irwig L, Gengos D. N-OF-1 trials comparing paracetamol and a nonsteroidal anti-inflammatory drug (NSAID) in osteoarthritis (OA). Australian and New Zealand Journal of Medicine. 1993; 23:537

129. Matsumoto AK, Babul N, Ahdieh H. Oxymorphone extended-release tablets relieve moderate to severe pain and improve physical function in osteoarthritis: results of a randomized, double-blind, placebo- and active-controlled phase III trial. Pain Medicine. 2005; 6(5):357-366

130. McIlwain H, Ahdieh H. Safety, tolerability, and effectiveness of oxymorphone extended release for moderate to severe osteoarthritis pain: a one-year study. American Journal of Therapeutics. 2005; 12(2):106-112

131. McKenna F, Borenstein D, Wendt H, Wallemark C, Lefkowith JB, Geis GS. Celecoxib versus diclofenac in the management of osteoarthritis of the knee. Scandinavian Journal of Rheumatology. 2001; 30(1):11-18

132. McKenna F, Weaver A, Fiechtner JJ, Bello AE, Fort JG. COX-2 specific inhibitors in the management of osteoarthritis of the knee: a placebo-controlled, randomized, double-blind study. JCR: Journal of Clinical Rheumatology. 2001; 7(3):151-159

133. Melo Gomes JA, Roth SH, Zeeh J, Bruyn GA, Woods EM, Geis GS. Double-blind comparison of efficacy and gastroduodenal safety of diclofenac/misoprostol, piroxicam, and naproxen in the treatment of osteoarthritis. Annals of the Rheumatic Diseases. 1993; 52(12):881-885

134. Miceli-Richard C, Le Bars M, Schmidely N, Dougados M. Paracetamol in osteoarthritis of the knee. Annals of the Rheumatic Diseases. 2004; 63(8):923-930

135. Moss P, Benson HAE, Will R, Wright A. Fourteen days of etoricoxib 60 mg improves pain, hyperalgesia and physical function in individuals with knee osteoarthritis: a randomized controlled trial. Osteoarthritis and Cartilage. 2017; 25(11):1781-1791

136. Muller-Fassbender H, Bach GL, Haase W, Rovati LC, Setnikar I. Glucosamine sulfate compared to ibuprofen in osteoarthritis of the knee. Osteoarthritis and Cartilage. 1994; 2(1):61-69

137. Munera C, Drehobl M, Sessler NE, Landau C. A randomized, placebo-controlled, double-blinded, parallel-group, 5-week study of buprenorphine transdermal system in adults with osteoarthritis. Journal of Opioid Management. 2010; 6(3):193-202

138. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated October 2020]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

139. Niethard FU, Gold MS, Solomon GS, Liu JM, Unkauf M, Albrecht HH et al. Efficacy of topical diclofenac diethylamine gel in osteoarthritis of the knee. Journal of Rheumatology. 2005; 32(12):2384-2392

140. Noack W, Fischer M, Forster KK, Rovati LC, Setnikar I. Glucosamine sulfate in osteoarthritis of the knee. Osteoarthritis and Cartilage. 1994; 2(1):51-59

141. Nowlan C, Wetmore S. Short report: ibuprofen versus glucosamine sulfate. Treating osteoarthritis pain. Canadian Family Physician. 2003; 49:1632-1634

142. Paul S, Das N, Ghosh S. The effects of aceclofenac and nabumetone in osteoarthritis. Jnma, Journal of the Nepal Medical Association. 2009; 48(174):121-125

143. Pavelka K, Gatterova J, Olejarova M, Machacek S, Giacovelli G, Rovati LC. Glucosamine sulfate use and delay of progression of knee osteoarthritis: a 3-year, randomized, placebo-controlled, double-blind study. Archives of Internal Medicine. 2002; 162(18):2113-2123

144. Pavelka K, Peliskova Z, Stehlikova H, Ratcliffe S, Repas C. Intraindividual differences in pain relief and functional improvement in osteoarthritis with diclofenac or tramadol. Clinical Drug Investigation. 1998; 16(6):421-429

145. Peloso PM, Bellamy N, Bensen W, Thomson GT, Harsanyi Z, Babul N et al. Double blind randomized placebo control trial of controlled release codeine in the treatment of osteoarthritis of the hip or knee. Journal of Rheumatology. 2000; 27(3):764-771

146. Peniston JH, Gold MS, Alwine LK. An open-label, long-term safety and tolerability trial of diclofenac sodium 1% gel in patients with knee osteoarthritis. Physician & Sportsmedicine. 2011; 39(3):31-38

147. Persson MSM, Stocks J, Sarmanova A, Fernandes G, Walsh DA, Doherty M et al. Individual responses to topical ibuprofen gel or capsaicin cream for painful knee osteoarthritis: a series of n-of-1 trials. Rheumatology. 2021; 60(5):2231-2237

148. Pincus T, Koch G, Lei H, Mangal B, Sokka T, Moskowitz R et al. Patient Preference for Placebo, Acetaminophen (paracetamol) or Celecoxib Efficacy Studies (PACES): two randomised, double blind, placebo controlled, crossover clinical trials in patients with knee or hip osteoarthritis. Annals of the Rheumatic Diseases. 2004; 63(8):931-939

149. Pincus T, Koch GG, Sokka T, Lefkowith J, Wolfe F, Jordan JM et al. A randomized, double-blind, crossover clinical trial of diclofenac plus misoprostol versus acetaminophen in patients with osteoarthritis of the hip or knee. Arthritis and Rheumatism. 2001; 44(7):1587-1598

150. Prior MJ, Harrison DD, Frustaci ME. A randomized, double-blind, placebo-controlled 12 week trial of acetaminophen extended release for the treatment of signs and symptoms of osteoarthritis. Current Medical Research and Opinion. 2014; 30(11):2377-2387

151. Pujalte JM, Llavore EP, Ylescupidez FR. Double-blind clinical evaluation of oral glucosamine sulphate in the basic treatment of osteoarthrosis. Current Medical Research and Opinion. 1980; 7(2):110-114

152. Puopolo A, Boice JA, Fidelholtz JL, Littlejohn TW, Miranda P, Berrocal A et al. A randomized placebo-controlled trial comparing the efficacy of etoricoxib 30 mg and ibuprofen 2400 mg for the treatment of patients with osteoarthritis. Osteoarthritis and Cartilage. 2007; 15(12):1348-1356

153. Qiu GX, Gao SN, Giacovelli G, Rovati L, Setnikar I. Efficacy and safety of glucosamine sulfate versus ibuprofen in patients with knee osteoarthritis. Arzneimittel-Forschung. 1998; 48(5):469-474

154. Reed K, Collaku A, Moreira S. Efficacy and safety of twice daily sustained-release paracetamol formulation for osteoarthritis pain of the knee or hip: a randomized, double-blind, placebo-controlled, twelve-week study. Current Medical Research and Opinion. 2018; 34(4):689-699

155. Rindone JP, Hiller D, Collacott E, Nordhaugen N, Arriola G. Randomized, controlled trial of glucosamine for treating osteoarthritis of the knee. Western Journal of Medicine. 2000; 172(2):91-94

156. Roth SH, Fuller P. Diclofenac topical solution compared with oral diclofenac: a pooled safety analysis. Journal of Pain Research. 2011; 4:159-167

157. Roth SH, Shainhouse JZ. Efficacy and safety of a topical diclofenac solution (pennsaid) in the treatment of primary osteoarthritis of the knee: a randomized, double-blind, vehicle-controlled clinical trial. Archives of Internal Medicine. 2004; 164(18):2017-2023

158. Rother M, Conaghan PG. A randomized, double-blind, phase III trial in moderate osteoarthritis knee pain comparing topical ketoprofen gel with ketoprofen-free gel. Journal of Rheumatology. 2013; 40(10):1742-1748

159. Rother M, Lavins BJ, Kneer W, Lehnhardt K, Seidel EJ, Mazgareanu S. Efficacy and safety of epicutaneous ketoprofen in Transfersome (IDEA-033) versus oral celecoxib

and placebo in osteoarthritis of the knee: multicentre randomised controlled trial. Annals of the Rheumatic Diseases. 2007; 66(9):1178-1183

160. Rovensky J, Micekova D, Gubzova Z, Fimmers R, Lenhard G, Vogtle-Junkert U et al. Treatment of knee osteoarthritis with a topical non-steroidal antiinflammatory drug. Results of a randomized, double-blind, placebo-controlled study on the efficacy and safety of a 5% ibuprofen cream. Drugs Under Experimental and Clinical Research. 2001; 27(5-6):209-221

161. Rozendaal RM, Koes BW, van Osch GJ, Uitterlinden EJ, Garling EH, Willemsen SP et al. Effect of glucosamine sulfate on hip osteoarthritis: a randomized trial. Annals of Internal Medicine. 2008; 148(4):268-277

162. Rozendaal RM, Uitterlinden EJ, van Osch GJ, Garling EH, Willemsen SP, Ginai AZ et al. Effect of glucosamine sulphate on joint space narrowing, pain and function in patients with hip osteoarthritis; subgroup analyses of a randomized controlled trial. Osteoarthritis and Cartilage. 2009; 17(4):427-432

163. Sanda M, Collins SH, Mahady J. Three-month multicenter study of etodolac (Ultradol) in patients with osteoarthritis of the hip. Current therapeutic research - clinical and experimental. 1983; 33(5):782-792

164. Sandelin J, Harilainen A, Crone H, Hamberg P, Forsskahl B, Tamelander G. Local NSAID gel (eltenac) in the treatment of osteoarthritis of the knee. A double blind study comparing eltenac with oral diclofenac and placebo gel. Scandinavian Journal of Rheumatology. 1997; 26(4):287-292

165. Sawitzke AD, Shi H, Finco MF, Dunlop DD, Bingham CO, 3rd, Harris CL et al. The effect of glucosamine and/or chondroitin sulfate on the progression of knee osteoarthritis: a report from the glucosamine/chondroitin arthritis intervention trial. Arthritis and Rheumatism. 2008; 58(10):3183-3191

166. Sawitzke AD, Shi H, Finco MF, Dunlop DD, Harris CL, Singer NG et al. Clinical efficacy and safety of glucosamine, chondroitin sulphate, their combination, celecoxib or placebo taken to treat osteoarthritis of the knee: 2-year results from GAIT. Annals of the Rheumatic Diseases. 2010; 69(8):1459-1464

167. Schiff MH. A comparison of Naprelan and Naprosyn in the treatment of osteoarthritis of the knee. American Journal of Orthopedics (Chatham, Nj). 1996; 25(9 Suppl):14-20

168. Schmitt W, Walter K, Kurth HJ. Clinical trial on the efficacy and safety of different diclofenac formulations: multiple-unit formulations compared to enteric coated tablets in patients with activated osteoarthritis. Inflammopharmacology. 1999; 7(4):363-375

169. Schnitzer T, Morton C, Coker S. Topical capsaicin therapy for osteoarthritis pain: achieving a maintenance regimen. Seminars in Arthritis and Rheumatism. 1994; 23(SUPPL. 3):34-40

170. Schnitzer TJ, Beier J, Geusens P, Hasler P, Patel SK, Senftleber I et al. Efficacy and safety of four doses of lumiracoxib versus diclofenac in patients with knee or hip primary osteoarthritis: a phase II, four-week, multicenter, randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism. 2004; 51(4):549-557

171. Schnitzer TJ, Dattani ID, Seriolo B, Schneider H, Moore A, Tseng L et al. A 13-week, multicenter, randomized, double-blind study of lumiracoxib in hip osteoarthritis. Clinical Rheumatology. 2011; 30(11):1433-1446

172. Schnitzer TJ, Hochberg MC, Marrero CE, Duquesroix B, Frayssinet H, Beekman M. Efficacy and safety of naproxcinod in patients with osteoarthritis of the knee: a 53-

week prospective randomized multicenter study. Seminars in Arthritis and Rheumatism. 2011; 40(4):285-297

173. Schnitzer TJ, Kivitz A, Frayssinet H, Duquesroix B. Efficacy and safety of naproxcinod in the treatment of patients with osteoarthritis of the knee: a 13-week prospective, randomized, multicenter study. Osteoarthritis and Cartilage. 2010; 18(5):629-639

174. Schnitzer TJ, Weaver AL, Polis AB, Petruschke RA, Geba GP, Vact et al. Efficacy of rofecoxib, celecoxib, and acetaminophen in patients with osteoarthritis of the knee. A combined analysis of the VACT studies. Journal of Rheumatology. 2005; 32(6):1093-1105

175. Schubiger BI, Ciccolunghi SN, Tanner K. Once daily dose treatment with a non-steroidal anti-rheumatic drug (diclofenac) in osteoarthrosis. Journal of International Medical Research. 1980; 8(2):167-174

176. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology. 2000; 39(10):1095-1101

177. Serrie A, Lange B, Steup A. Tapentadol prolonged-release for moderate-to-severe chronic osteoarthritis knee pain: a double-blind, randomized, placebo- and oxycodone controlled release-controlled study. Current Medical Research and Opinion. 2017; 33(8):1423-1432

178. Sheldon E, Beaulieu A, Paster Z, Dutta D, Yu S, Sloan VS. Efficacy and tolerability of lumiracoxib in the treatment of osteoarthritis of the knee: a 13-week, randomized, double-blind comparison with celecoxib and placebo. Clinical Therapeutics. 2005; 27(1):64-77

179. Sheldon EA, Beaulieu A, Paster Z, Yu S, Rebuli R. Long-term efficacy and safety of lumiracoxib 100 mg: an open-label extension of a 13-week randomized controlled trial in patients with primary osteoarthritis of the knee. Clinical and Experimental Rheumatology. 2008; 26(4):611-619

180. Sikes DH, Agrawal NM, Zhao WW, Kent JD, Recker DP, Verburg KM. Incidence of gastroduodenal ulcers associated with valdecoxib compared with that of ibuprofen and diclofenac in patients with osteoarthritis. European Journal of Gastroenterology and Hepatology. 2002; 14(10):1101-1111

181. Simon LS, Grierson LM, Naseer Z, Bookman AA, Zev Shainhouse J. Efficacy and safety of topical diclofenac containing dimethyl sulfoxide (DMSO) compared with those of topical placebo, DMSO vehicle and oral diclofenac for knee osteoarthritis. Pain. 2009; 143(3):238-245

182. Smugar SS, Schnitzer TJ, Weaver AL, Rubin BR, Polis AB, Tershakovec AM. Rofecoxib 12.5 mg, rofecoxib 25 mg, and celecoxib 200 mg in the treatment of symptomatic osteoarthritis: results of two similarly designed studies. Current Medical Research and Opinion. 2006; 22(7):1353-1367

183. Sofat N, Harrison A, Russell MD, Ayis S, Kiely PD, Baker EH et al. The effect of pregabalin or duloxetine on arthritis pain: a clinical and mechanistic study in people with hand osteoarthritis. Journal of Pain Research. 2017; 10:2437-2449

184. Strand V, Bergman M, Singh JA, Gibofsky A, Kivitz A, Young C. Low-dose SoluMatrix diclofenac in patients with osteoarthritis pain: impact on quality of life in a controlled trial. Clinical Rheumatology. 2017; 36(6):1357-1367

185. Tannenbaum H, Berenbaum F, Reginster JY, Zacher J, Robinson J, Poor G et al. Lumiracoxib is effective in the treatment of osteoarthritis of the knee: a 13 week, randomised, double blind study versus placebo and celecoxib. Annals of the Rheumatic Diseases. 2004; 63(11):1419-1426

186. Temple AR, Benson GD, Zinsenheim JR, Schweinle JE. Multicenter, randomized, double-blind, active-controlled, parallel-group trial of the long-term (6-12 months) safety of acetaminophen in adult patients with osteoarthritis. Clinical Therapeutics. 2006; 28(2):222-235

187. Thorne C, Beaulieu AD, Callaghan DJ, O'Mahony WF, Bartlett JM, Knight R et al. A randomized, double-blind, crossover comparison of the efficacy and safety of oral controlled-release tramadol and placebo in patients with painful osteoarthritis. Pain Research & Management. 2008; 13(2):93-102

188. Tiso RL, Tong-Ngork S, Fredlund KL. Oral versus topical Ibuprofen for chronic knee pain: a prospective randomized pilot study. Pain Physician. 2010; 13(5):457-467

189. Trnavsky K, Fischer M, Vogtle-Junkert U, Schreyger F. Efficacy and safety of 5% ibuprofen cream treatment in knee osteoarthritis. Results of a randomized, double-blind, placebo-controlled study. Journal of Rheumatology. 2004; 31(3):565-572

190. Trudeau J, Van Inwegen R, Eaton T, Bhat G, Paillard F, Ng D et al. Assessment of pain and activity using an electronic pain diary and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee. Pain Practice. 2015; 15(3):247-255

191. Truitt KE, Sperling RS, Ettinger WH, Jr., Greenwald M, DeTora L, Zeng Q et al. A multicenter, randomized, controlled trial to evaluate the safety profile, tolerability, and efficacy of rofecoxib in advanced elderly patients with osteoarthritis. Aging-Clinical & Experimental Research. 2001; 13(2):112-121

192. Tugwell PS, Wells GA, Shainhouse JZ. Equivalence study of a topical diclofenac solution (pennsaid) compared with oral diclofenac in symptomatic treatment of osteoarthritis of the knee: a randomized controlled trial. Journal of Rheumatology. 2004; 31(10):2002-2012

193. Uchio Y, Enomoto H, Alev L, Kato Y, Ishihara H, Tsuji T et al. A randomized, double-blind, placebo-controlled Phase III trial of duloxetine in Japanese patients with knee pain due to osteoarthritis. Journal of Pain Research. 2018; 11:809-821

194. Uchio Y, Enomoto H, Ishida M, Tsuji T, Ochiai T, Konno S. Safety and efficacy of duloxetine in Japanese patients with chronic knee pain due to osteoarthritis: an open-label, long-term, Phase III extension study. Journal of Pain Research. 2018; 11:1391-1403

195. Underwood M, Ashby D, Carnes D, Castelnuovo E, Cross P, Harding G et al. Topical or oral ibuprofen for chronic knee pain in older people. The TOIB study. Health technology assessment (Winchester, England). 2008; 12(22):iii-iv, ix-155

196. Underwood M, Ashby D, Cross P, Hennessy E, Letley L, Martin J et al. Advice to use topical or oral ibuprofen for chronic knee pain in older people: randomised controlled trial and patient preference study. BMJ. 2008; 336(7636):138-142

197. Verkleij SP, Luijsterburg PA, Willemsen SP, Koes BW, Bohnen AM, Bierma-Zeinstra SM. Effectiveness of diclofenac versus paracetamol in knee osteoarthritis: a randomised controlled trial in primary care. British Journal of General Practice. 2015; 65(637):e530-537

198. Via RM. Diclofenac more effective than acetaminophen in osteoarthritis of the knee. Journal of Family Practice. 2003; 52(7):512-513

199. Vojtassak J, Vojtassak J, Jacobs A, Rynn L, Waechter S, Richarz U. A Phase IIIb, Multicentre, Randomised, Parallel-Group, Placebo-Controlled, Double-Blind Study to Investigate the Efficacy and Safety of OROS Hydromorphone in Subjects with Moderate-to-Severe Chronic Pain Induced by Osteoarthritis of the Hip or the Knee. Pain Research and Treatment. 2011; 2011:239501

200. Wadsworth LT, Kent JD, Holt RJ. Efficacy and safety of diclofenac sodium 2% topical solution for osteoarthritis of the knee: a randomized, double-blind, vehicle-controlled, 4 week study. Current Medical Research and Opinion. 2016; 32(2):241-250

201. Wang G, Bi L, Li X, Li Z, Zhao D, Chen J et al. Efficacy and safety of duloxetine in Chinese patients with chronic pain due to osteoarthritis: a randomized, double-blind, placebo-controlled study. Osteoarthritis and Cartilage. 2017; 25(6):832-838

202. Wanka J, Dixon A. Treatment of osteoarthritis of the hip with indomethacin: a controlled clinical trial. Annals of rheumatic diseases. 1964; 23:288-294

203. Wasserman A, Proctor J. Carprofen in osteoarthritis: a comparison with indomethacin and placebo. Advances in Therapy. 1984; 1(4):256-266

204. Wiesenhutter CW, Boice JA, Ko A, Sheldon EA, Murphy FT, Wittmer BA et al. Evaluation of the comparative efficacy of etoricoxib and ibuprofen for treatment of patients with osteoarthritis: A randomized, double-blind, placebo-controlled trial. Mayo Clinic Proceedings. 2005; 80(4):470-479

205. Williams GW, Ettlinger RE, Ruderman EM, Hubbard RC, Lonien ME, Yu SS et al. Treatment of osteoarthritis with a once-daily dosing regimen of celecoxib: a randomized, controlled trial. JCR: Journal of Clinical Rheumatology. 2000; 6(2):65-74

206. Williams GW, Hubbard RC, Yu SS, Zhao W, Geis GS. Comparison of once-daily and twice-daily administration of celecoxib for the treatment of osteoarthritis of the knee. Clinical Therapeutics. 2001; 23(2):213-227

207. Williams HJ, Ward JR, Egger MJ, Neuner R, Brooks RH, Clegg DO et al. Comparison of naproxen and acetaminophen in a two-year study of treatment of osteoarthritis of the knee. Arthritis and Rheumatism. 1993; 36(9):1196-1206

208. Wittenberg RH, Schell E, Krehan G, Maeumbaed R, Runge H, Schlüter P et al. First-dose analgesic effect of the cyclo-oxygenase-2 selective inhibitor lumiracoxib in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled comparison with celecoxib. Arthritis Research & Therapy. 2006; 8(2):R35

209. Yocum D, Fleischmann R, Dalgin P, Caldwell J, Hall D, Roszko P. Safety and efficacy of meloxicam in the treatment of osteoarthritis: a 12-week, double-blind, multiple-dose, placebo-controlled trial. The Meloxicam Osteoarthritis Investigators. Archives of Internal Medicine. 2000; 160(19):2947-2954

210. Zautra AJ, Smith BW. Impact of controlled-release oxycodone on efficacy beliefs and coping efforts among osteoarthritis patients with moderate to severe pain. Clinical Journal of Pain. 2005; 21(6):471-477

211. Zenk JL, Helmer TR, Kuskowski MA. The effects of milk protein concentrate on the symptoms of osteoarthritis in adults: an exploratory, randomized, double-blind, placebo-controlled trial. Current therapeutic research - clinical and experimental. 2002; 63(7):430-442

212. Zhao SZ, McMillen JI, Markenson JA, Dedhiya SD, Zhao WW, Osterhaus JT et al. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis treated with celecoxib. Pharmacotherapy:The Journal of Human Pharmacology & Drug Therapy. 1999; 19(11):1269-1278

Appendices

Appendix A – Review protocols

Review protocol for the clinical and cost-effectiveness of pharmacological interventions for the management of osteoarthritis

ID	Field	Content
0.	PROSPERO registration number	
1.	Review title	What is the clinical and cost-effectiveness of oral, topical and transdermal medicines for the management of osteoarthritis?
2.	Review question	4.1 What is the clinical and cost-effectiveness of oral, topical and transdermal medicines for the management of osteoarthritis?
3.	Objective	To evaluate the clinical and cost-effectiveness of oral, topical and transdermal pharmacological interventions in the management of osteoarthritis. A range of pharmacological interventions have been reported to reduce joint pain and improve function. However, these interventions are not used consistently
4.	Searches	The following databases will be searched (all years): • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE Searches will be restricted by: • English language • Human studies

		• Letters and comments are excluded Other searches: • Inclusion lists of relevant systematic reviews will be checked by the reviewer. The searches may be re-run 6 weeks before final submission of the review and further studies retrieved for inclusion if relevant. The full search strategies for MEDLINE database will be published in the final review.
5.	Condition or domain being studied	Osteoarthritis in adults (defined as a clinical diagnosis of osteoarthritis with or without imaging)
6.	Population	Inclusion: • Adults (age ≥16 years) with osteoarthritis affecting any joint Stratify by site of osteoarthritis for topical (local) interventions only: • Hip • Knee • Ankle • Foot • Toe • Shoulder • Elbow • Wrist • Hand • Thumb

		• Finger • Temporomandibular joint (TMJ) • Multisite To note that where evidence for other rare forms of osteoarthritis is identified the committee will stratify into a group they are most similar to. Exclusion: • Children (age <16 years) • People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, hemochromatosis, haemophilic arthropathy, diseases of childhood that may predispose to osteoarthritis, and malignancy). • Studies with an unclear population (e,g, proportion of participants with osteoarthritis unclear) • Spinal osteoarthritis
7.	Intervention	Oral medicines: • Paracetamol • Non-specific NSAIDs (including ibuprofen, diclofenac, naproxen, mefenamic acid, indomethacin and high-dose aspirin) • Non-specific NSAIDs with gastroprotection • Specific COX-2 inhibitors (including celecoxib, etodolac, meloxicam and) • Specific COX-2 inhibitors with gastroprotection • Weak opioids (including codeine and dihydrocodeine) • Strong opioids (including morphine, oxycodone, hydromorphone, tramadol and tapentadol) • Anti-epileptics (including gabapentin and pregabalin)

		• Antidepressants (including SSRIs, SNRIs and tricyclics) • Glucosamine Topical (local) medicines: • Capsaicin cream • NSAID gels (e.g. ibuprofen, diclofenac) • Rubefacients • Topical local anaesthetics Topical (systemic e.g. transdermal) medicines: • Opioids Note: only medicines licensed for use in the UK will be included.
8.	Comparator	• Each other (class comparisons) • Placebo
9.	Types of study to be included	• Systematic reviews of RCTs • Parallel RCTs • Cross-over RCTs will be considered if insufficient evidence is available from parallel RCTs* Non-randomised studies will be excluded.

		*Insufficient evidence defined as evidence that is insufficient to inform recommendations (either quality or quantity).
10.	Other exclusion criteria	• Non-English language studies • Non-randomised/observational studies Abstracts will be excluded as it is expected there will be sufficient full text published studies available.
11.	Context	N/A
12.	Primary outcomes (critical outcomes)	Stratify by ≤/>3 months (longest time-point in each): • Health-related quality of life [validated patient-reported outcomes, continuous data prioritised] • Pain [validated patient-reported outcomes, continuous data prioritised] • Physical function [validated patient-reported outcomes, continuous data prioritised]
13.	Secondary outcomes (important outcomes)	• Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flares [validated patient-reported outcomes, continuous data prioritised] • Serious adverse events 1A: Gastrointestinal (bleeding or perforation) adverse events • Serious adverse events 1B: Gastrointestinal (non-bleeding or perforation) adverse events • Serious adverse events 2: Cardiovascular adverse events • Serious adverse events 3: Hepatorenal adverse events • Serious adverse events 4: Central nervous system adverse events
14.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third

		independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. EviBASE will be used for data extraction. Study investigators may be contacted for missing data where time and resources allow.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual For intervention reviews the following checklists will be used according to the study design being assessed: • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
16.	Strategy for data synthesis	• Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5). • GRADEpro will be used to assess the quality of evidence for each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision)

		will be appraised for each outcome. Publication bias is tested for when there are more than 5 studies for an outcome. The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ • Where meta-analysis is not possible, data will be presented and quality assessed individually per outcome. • WinBUGS will be used for network meta-analysis, if possible given the data identified. Heterogeneity between studies in the effect measures will be assessed using the I^2 statistic and visual inspection. We will consider an I^2 value great than 50% as indicative of substantial heterogeneity. If significant heterogeneity is identified during meta-analysis then subgroup analysis, using subgroups predefined by the GC, will take place. If this does not explain the heterogeneity, the results will be presented using a random-effects model.
17.	Analysis of sub-groups	Analyses to be conducted if heterogeneity in the meta-analysis is present: Two-step sensitivity analysis: 1: Exclude all enrichment studies* 2: Exclude only the most selective enrichment style studies (those including only known responders) Subgroup analyses: • Diagnosis with or without imaging (indicative of severity) • Multimorbidity (high versus low morbidity score; as defined by study, measured by validated instruments e.g. Charlson Comorbidity Index) • Age (≤/> 75 years) • Site of osteoarthritis (for systemic treatments only)

<table>
<tr><td></td><td></td><td colspan="3">• For studies where the intervention or comparison is glucosamine only: statement of quality assurance of glucosamine product

*Enrichment studies defined as those that select a population based on previous response to treatment (e.g. including only those known to respond to a particular medicine)</td></tr>
<tr><td rowspan="8">18.</td><td rowspan="8">Type and method of review</td><td>☒</td><td colspan="2">Intervention</td></tr>
<tr><td>☐</td><td colspan="2">Diagnostic</td></tr>
<tr><td>☐</td><td colspan="2">Prognostic</td></tr>
<tr><td>☐</td><td colspan="2">Qualitative</td></tr>
<tr><td>☐</td><td colspan="2">Epidemiologic</td></tr>
<tr><td>☐</td><td colspan="2">Service Delivery</td></tr>
<tr><td>☐</td><td colspan="2">Other (please specify)</td></tr>
<tr><td></td><td></td><td colspan="3"></td></tr>
<tr><td>19.</td><td>Language</td><td colspan="3">English</td></tr>
<tr><td>20.</td><td>Country</td><td colspan="3">England</td></tr>
<tr><td>21.</td><td>Anticipated or actual start date</td><td colspan="3">23/08/2019</td></tr>
<tr><td>22.</td><td>Anticipated completion date</td><td colspan="3">25/08/2021</td></tr>
<tr><td rowspan="3">23.</td><td rowspan="3">Stage of review at time of this submission</td><td>Review stage</td><td>Started</td><td>Completed</td></tr>
<tr><td>Preliminary searches</td><td>☑</td><td>☐</td></tr>
<tr><td>Piloting of the study selection process</td><td>☐</td><td>☐</td></tr>
</table>

		Formal screening of search results against eligibility criteria	☐	☐
		Data extraction	☐	☐
		Risk of bias (quality) assessment	☐	☐
		Data analysis	☐	☐
24.	Named contact	5a. Named contact National Guideline Centre 5b Named contact e-mail [Guideline email]@nice.org.uk [Developer to check with Guideline Coordinator for email address] 5e Organisational affiliation of the review National Institute for Health and Care Excellence (NICE) and the National Guideline Centre		
25.	Review team members	From the National Guideline Centre: Carlos Sharpin [Guideline lead] Julie Neilson [Senior systematic reviewer] George Wood [Systematic reviewer] Emma Cowles [Senior health economist] Joseph Runicles [Information specialist] Amber Hernaman [Project manager]		

26.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10127
29.	Other registration details	TBC
30.	Reference/URL for published protocol	TBC
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.

32.	Keywords	Adults; COX-2 inhibitors; Intervention; Neuropathic pain; NSAIDs; Opioids; Oral; Osteoarthritis; Paracetamol; Pharmacological; Topical; Transdermal	
33.	Details of existing review of same topic by same authors	None	
34.	Current review status	☒	Ongoing
		☐	Completed but not published
		☐	Completed and published
		☐	Completed, published and being updated
		☐	Discontinued
35..	Additional information	N/A	
36.	Details of final publication	www.nice.org.uk	

1
2

Table 1. Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken for all years using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2005, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Studies published in 2005 or later, that were included in the previous guidelines, will be reassessed for inclusion and may be included or selectively excluded based on their relevance to the questions covered in this update and whether more applicable evidence is also identified. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[138] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included.

Where there is discretion

The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below.

The health economist will be guided by the following hierarchies.

Setting:

- UK NHS (most applicable).
- OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden).
- OECD countries with predominantly private health insurance systems (for example, Switzerland).
- Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations.

Health economic study type:

- Cost–utility analysis (most applicable).
- Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis).
- Comparative cost analysis.
- Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations.

Year of analysis:

- The more recent the study, the more applicable it will be.
- Studies published in 2005 or later (including any such studies included in the previous guidelines) but that depend on unit costs and resource data entirely or predominantly from before 2005 will be rated as 'Not applicable'.
- Studies published before 2005 (including any such studies included in the previous guidelines) will be excluded before being assessed for applicability and methodological limitations.

Quality and relevance of effectiveness data used in the health economic analysis:

- The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

1

Appendix B – Literature search strategies

- What is the clinical and cost-effectiveness of oral, topical and transdermal medicines for the management of osteoarthritis?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[138]

For more information, please see the Methodology review published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using an Osteoarthritis population. All results were then sifted for each question. Search filters were applied to the search where appropriate.

Table 2: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 17 November 2021	Randomised controlled trials Systematic review studies Exclusions (animals studies, letters, comments)
Embase (OVID)	1974 – 17 November 2021	Randomised controlled trials Systematic review studies Exclusions (animals studies, letters, comments)
The Cochrane Library (Wiley)	Cochrane Reviews to 2021 Issue 11 of 12 CENTRAL to 2021 Issue 11 of 12	None

Medline (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.

15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice or rodent*).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	randomized controlled trial.pt.
28.	controlled clinical trial.pt.
29.	randomi#ed.ti,ab.
30.	placebo.ab.
31.	randomly.ti,ab.
32.	Clinical Trials as topic.sh.
33.	trial.ti.
34.	or/27-33
35.	Meta-Analysis/
36.	exp Meta-Analysis as Topic/
37.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
38.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
39.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
40.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
41.	(search* adj4 literature).ab.
42.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
43.	cochrane.jw.
44.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
45.	or/35-44
46.	26 and (34 or 45)

Embase (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/

11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice or rodent*).ti.
22.	or/14-21
23.	6 not 22
24.	Limit 23 not English language
25.	random*.ti,ab.
26.	factorial*.ti,ab.
27.	(crossover* or cross over*).ti,ab.
28.	((doubl* or singl*) adj blind*).ti,ab.
29.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
30.	crossover procedure/
31.	single blind procedure/
32.	randomized controlled trial/
33.	double blind procedure/
34.	or/25-33
35.	systematic review/
36.	meta-analysis/
37.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
38.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
39.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
40.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
41.	(search* adj4 literature).ab.
42.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
43.	cochrane.jw.
44.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
45.	or/35-44
46.	24 and (34 or 45)

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Osteoarthritis] explode all trees
#2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*):ti,ab
#3.	(degenerative near/2 arthritis):ti,ab
#4.	coxarthrosis:ti,ab

#5.	gonarthrosis:ti,ab
#6.	(or #1-#5)

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to a Gout population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA – this ceased to be updates after March 2018). NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics studies and quality of life studies. Searches for quality of life studies were run for general information.

Table 3: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	1 January 2014 – 17 November 2021	Health economics studies Quality of life studies Exclusions (animals studies, letters, comments)
Embase	1 January 2014 – 17 November 2021	Health economics studies Quality of life studies Exclusions (animals studies, letters, comments)
Centre for Research and Dissemination (CRD)	HTA - Inception – 31 March 2018 NHSEED - Inception to 31 March 2015	None

Medline (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.

17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice or rodent*).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	Economics/
28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/
31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/27-42
44.	quality-adjusted life years/
45.	sickness impact profile/
46.	(quality adj2 (wellbeing or well being)).ti,ab.
47.	sickness impact profile.ti,ab.
48.	disability adjusted life.ti,ab.
49.	(qal* or qtime* or qwb* or daly*).ti,ab.
50.	(euroqol* or eq5d* or eq 5*).ti,ab.
51.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
52.	(hui or hui1 or hui2 or hui3).ti,ab.
53.	(health* year* equivalent* or hye or hyes).ti,ab.
54.	discrete choice*.ti,ab.
55.	rosser.ti,ab.

56.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
57.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
58.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
59.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
60.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
61.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
62.	or/44-61
63.	26 and (43 or 62)

Embase (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice or rodent*).ti.
22.	or/14-21
23.	6 not 22
24.	Limit 23 to English language
25.	health economics/
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/

31.	budget*.ti,ab.
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	quality adjusted life year/
40.	"quality of life index"/
41.	short form 12/ or short form 20/ or short form 36/ or short form 8/
42.	sickness impact profile/
43.	(quality adj2 (wellbeing or well being)).ti,ab.
44.	sickness impact profile.ti,ab.
45.	disability adjusted life.ti,ab.
46.	(qal* or qtime* or qwb* or daly*).ti,ab.
47.	(euroqol* or eq5d* or eq 5*).ti,ab.
48.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
49.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
50.	(hui or hui1 or hui2 or hui3).ti,ab.
51.	(health* year* equivalent* or hye or hyes).ti,ab.
52.	discrete choice*.ti,ab.
53.	rosser.ti,ab.
54.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
55.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
56.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
57.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
58.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
59.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
60.	or/39-59
61.	24 and (38 or 60)

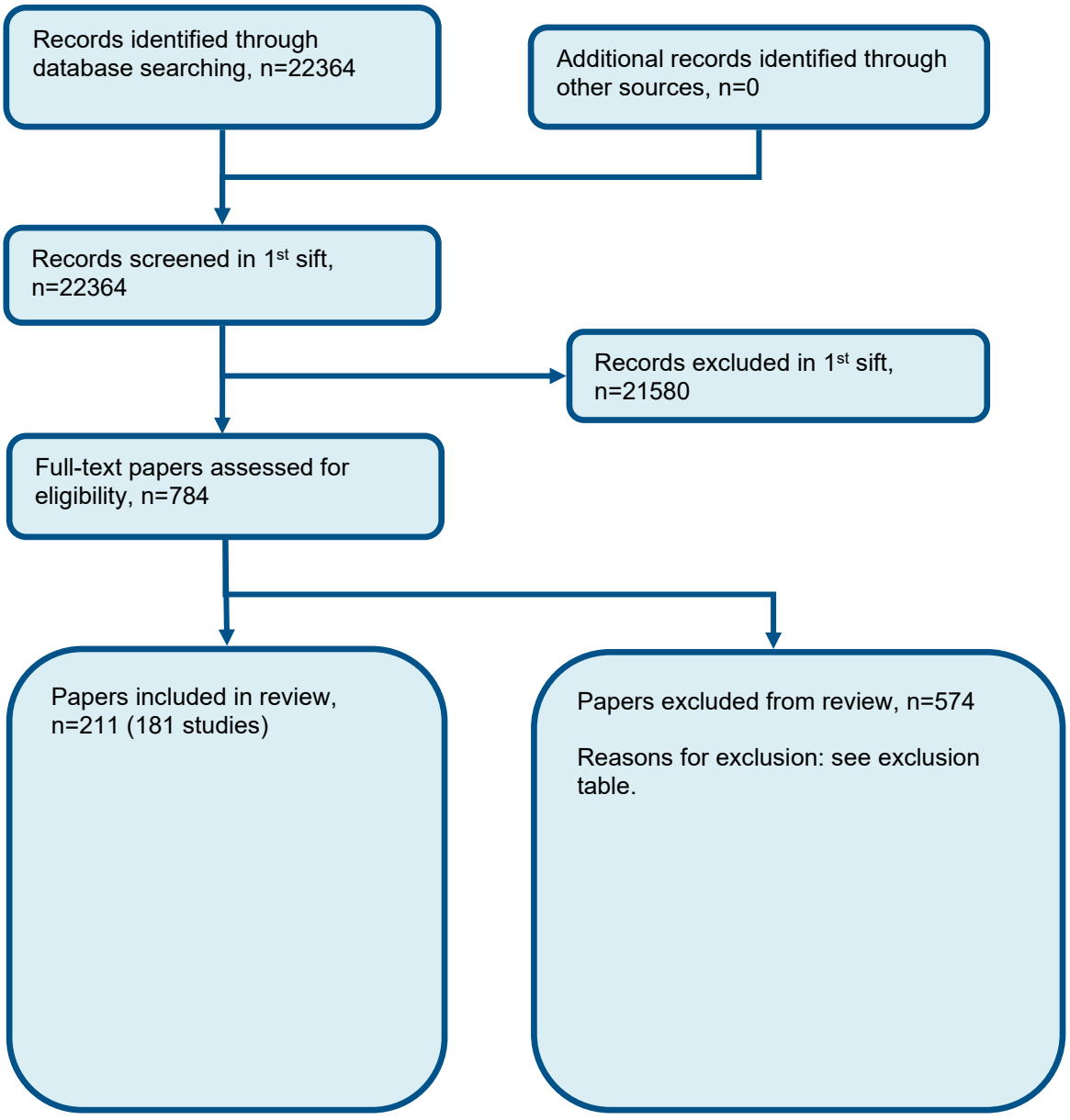
NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Osteoarthritis EXPLODE ALL TREES
#2.	((osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*))
#3.	((degenerative adj2 arthritis))
#4.	(coxarthrosis)
#5.	(gonarthrosis)
#6.	#1 OR #2 OR #3 OR #4 OR #5
#7.	(#6) IN NHSEED

#8.	(#6) IN HTA

Appendix C – Effectiveness evidence study selection

Figure 1: Flow chart of clinical study selection for the review of the clinical and cost-effectiveness of pharmacological interventions for the management of osteoarthritis



Appendix D – Effectiveness evidence

Study	Abou-raya 2012[1]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=288)
Countries and setting	Conducted in Egypt; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 16 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: American College of Rheumatology clinical and radiographic criteria of primary knee osteoarthritis, with knee pain (>40 on the 24 hour average pain severity scale) using mean of daily ratings from the week preceding randomisation for >14 days/month during three consecutive months prior to enrolment. Kellgren Lawrence grade I-III tibiofemoral or patellofemoral osteoarthritis on weight bearing anteroposterior and sunrise view radiographs.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with American College of Rheumatology clinical and radiographic criteria of primary knee osteoarthritis, with knee pain (>40 on the 24 hour average pain severity scale) using mean of daily ratings from the week preceding randomisation for >14 days/month during three consecutive months prior to enrolment. Kellgren Lawrence grade I-III tibiofemoral or patellofemoral osteoarthritis on weight bearing anteroposterior and sunrise view radiographs.
Exclusion criteria	Morbid obesity (BMI greater than 32 kg/m²); joint inflammatory diseases and/or crystal-induced arthropathies; any other concomitant disease (such as neuropsychiatric disease including cognitive impairment, Alzheimer's disease, Parkinson's disease, cerebrovascular disease, cardiovascular disease, liver and renal disease) or were taking any other antidepressants that could interfere with the evaluation of the intervention
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 68.7 (6.0). Gender (M:F): 47:241. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Low comorbidity score (184 had 0-1 comorbidities. 38 had at least 2 comorbidities.). 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Kellgren Lawrence grade II-III (majority grade II) Duration of symptoms (mean [SD]): 5.7 (4.7) years
Indirectness of population	No indirectness
Interventions	(n=144) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 60mg/day. Duration 16 weeks. Concurrent medication/care: Concomitant rescue medication use, including paracetamol up to 4g/day and NSAIDs was allowed to continue provided they did not increase the dose. Indirectness: No indirectness (n=144) Intervention 2: Placebo. Matching placebo. Duration 16 weeks. Concurrent medication/care: Concomitant rescue medication use, including paracetamol up to 4g/day and NSAIDs was allowed to continue provided they did not increase the dose. Indirectness: No indirectness
Funding	Equipment / drugs provided by industry (The authors acknowledge the supported of the pharmaceutical industry, Eli Lilly, Egypt, for providing the medicines for the study. The industry had no role in approval or preparation of the published manuscript.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain score at 16 weeks; Group 1: mean 6 (SD 4.1); n=144, Group 2: mean 8.4 (SD 5.4); n=144; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline duloxetine: 9.1 (4.6). Baseline placebo: 8.9 (5.1).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI, comorbidities, analgesic use, paracetamol use, disease duration, baseline values for outcomes, Kellgren-Lawrence grade and activities of daily living; Group 1 Number missing: 20, Reason: 6 lost to follow up. 9 adverse events. 5 lack of efficacy.; Group 2 Number missing: 13, Reason: 5 lost to follow up. 6 adverse events. 2 lack of efficacy.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function score at 16 weeks; Group 1: mean 24.6 (SD 8.4); n=144, Group 2: mean 30.3 (SD 9.8); n=144; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline duloxetine: 33.1 (7.5). Baseline placebo: 33.5 (7.1).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI, comorbidities, analgesic use, paracetamol use, disease duration, baseline values for outcomes, Kellgren-Lawrence grade and activities of daily living; Group 1 Number missing: 20, Reason: 6 lost to follow up. 9 adverse events. 5 lack of efficacy.; Group 2 Number missing: 13, Reason: 5 lost to follow up. 6 adverse events. 2 lack of efficacy.

Protocol outcome 3: Psychological distress at ≤3- or >3- months
- Actual outcome for Knee: Geriatric Depression Scale score at 16 weeks; Group 1: mean 5.2 (SD 1.7); n=144, Group 2: mean 9.7 (SD 2.2); n=144; Geriatric

Depression Scale 0-15 Top=High is poor outcome; Comments: Baseline duloxetine: 9.9 (2.7). Baseline placebo: 9.8 (2.5). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI, comorbidities, analgesic use, paracetamol use, disease duration, baseline values for outcomes, Kellgren-Lawrence grade and activities of daily living; Group 1 Number missing: 20, Reason: 6 lost to follow up. 9 adverse events. 5 lack of efficacy.; Group 2 Number missing: 13, Reason: 5 lost to follow up. 6 adverse events. 2 lack of efficacy.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	ACTA trial: Pincus 2001[149]
Study type	RCT (Patient randomised; Crossover: 3-7 days)
Number of studies (number of participants)	1 (n=227)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks for each intervention arm, two 3-7 day washout periods
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the hip or knee with Kellgren-Lawrence radiographic grade 2-4 changes and a visual analogue pain score ≥30mm
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Age >40 years; Kellgren Lawrence radiographic grade 2-4 osteoarthritis of the hip or knee; visual analogue pain scale score of ≥30mm (range 0-100mm).
Exclusion criteria	Severe comorbidities; hypersensitivity to paracetamol, diclofenac or misoprostol
Recruitment/selection of patients	227 people were enrolled (multisite study) of those 218 provided data for the first treatment period. People were given the choice as to whether they wanted to be involved in the second treatment period. 181 provided data for both treatment periods.
Age, gender and ethnicity	Age - Other: Mean (SE): 61.5 (1.37). Gender (M:F): 86:141. Ethnicity: 177 were Caucasian. They did not specify the ethnicity of other participants.
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip - the number of people with each was not stated).
Extra comments	Severity: Kellgren Lawrence radiographic grade (mean [SE]): 2.8 (0.089). Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=227) Intervention 1: NSAID and gastroprotection - NSAID and misoprostol . 75mg diclofenac with 200 micrograms misoprostol to be taken twice a day (oral). Placebo pills were taken (to make it so tablets were being taken four times a day) to blind people to allocation. Duration 6 weeks. Concurrent medication/care: Propoxyphene could be taken as rescue medication during the washout period. Indirectness: No indirectness

	(n=227) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg four times a day (oral). Duration 6 weeks. Concurrent medication/care: Propoxyphene could be taken as rescue medication during the washout period. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Pharmacia)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NSAID AND MISOPROSTOL versus PARACETEMOL

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 bodily pain subscale at 6 weeks; MD; 3.83 (SE: 0.75) SF-36 bodily pain subscale 0-100 Top=High is good outcome, Comments: P value = <0.001;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: MDHAQ (multidimensional health assessment questionnaire) visual analogue pain scale at 6 weeks; MD; -14.6 (SE: 1.81) MDHAQ visual analogue pain scale 0-100 Top=High is poor outcome, Comments: P = <0.001;
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the

diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Haemorrhagic diarrhoea at 6 weeks; Group 1: 1/218, Group 2: 0/218
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Myocardial infarction at 6 weeks; Group 1: 2/218, Group 2: 1/218
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).

Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Total with abnormal SGOT level at 6 weeks; Group 1: 22/218, Group 2: 10/218
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period.

2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2). Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 6 weeks; Group 1: 5/218, Group 2: 7/218 Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, radiographic grade and features, and outcome baseline values; Group 1 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).; Group 2 Number missing: 120, Reason: 227 people randomised. First arm: 72 completed period 1 in the diclofenac and misoprostal arm (6 had no data after visit 2, 18 early terminations, 3 late visits, 13 six week visits with poor compliance, 12 adverse events, 3 inefficacy). 26 of those people continued for period 2. 3 people chose not to continue. After this 95 joined the second period. 2 had no data after visit 4. There were 20 early terminations, 5 adverse events, 13 inefficacy, 2 others, 1 late visit, and 10 sixth week visits with poor compliance. 62 completed period 2 per protocol (with 51 completing both period 1 and 2). Second arm: 82 completed period 1 per protocol. 115 were given paracetamol. 2 had no data after visit 2, 1 missed visit 3 but continued, there were 23 early terminations, 6 adverse events, 11 inefficacy, 6 others, 1 late visit, and 6 sixth week visit with poor compliance. 7 decided to not continue. 16 of the people who discontinued continued to period 2. 91 were given diclofenac and misoprostol. 2 had no data after visit 4. There were 16 early terminations, 11 for adverse events, 2 for inefficacy, 3 others. There were 2 late visits and 5 sixth-week visits with poor compliance. 66 complete period 2 per protocol (with 69 completing period 1 and 2).	
Protocol outcomes not reported by the study	Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Afilalo 2010[2]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1023)
Countries and setting	Conducted in Australia, Canada, New Zealand, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (with an additional 3 week titration period)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of knee osteoarthritis according to the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women ≥40 years of age with a diagnosis of osteoarthritis of the knee according to the American College of Rheumatology criteria, functional capacity class I-III, and pain at the reference joint requiring the use of analgesics (non-opioids or opioids at doses equivalent to ≤160mg oral morphine/day) for ≥3 months prior to screening. People were dissatisfied with their current analgesic therapy and had an average baseline pain intensity numerical rating scale score of ≥5 during the 3 days preceding randomisation
Exclusion criteria	The presence of clinically significant or unstable medical or psychiatric disease; requirement for painful procedures (e.g. surgery) during the study that could influence efficacy or safety assessments; a history of substance abuse, epilepsy/seizure disorder, stroke/transient ischaemic attack, malignancy (preceding 2 years), HIV infection, chronic hepatitis B or C, uncontrolled hypertension (systolic blood pressure >160mmHg and/or diastolic blood pressure >95mmHg), severe renal impairment (creatinine clearance <60mL/min), moderate or severe hepatic impairment, ALT or AST concentrations >3 times the upper limit of normal, and hypersensitivity to study medications or their excipients; people with conditions potentially influencing the assessment of osteoarthritis pain (anatomical deformities, fibromyalgia, gout or infectious or autoimmune diseases affecting the knee; the use of concomitant analgesics (except allowed doses of paracetamol); neuroleptics, tricyclic antidepressants, anticonvulsants, antiparkinsonian drugs and serotonin-norepinephrine reuptake inhibitors within 14 days prior to screening and during the study; other medication use that has not been at a stable dose for ≥3 months; corticosteroids use within 4 weeks to 6 months prior to screening, depending on the route of administration

Recruitment/selection of patients	People were recruited from 87 sites in the US, 15 sites in Canada, 6 sites in New Zealand, and 4 sites in Australia
Age, gender and ethnicity	Age - Mean (SD): 58.3 (9.9). Gender (M:F): 405:618. Ethnicity: White = 772, Black = 132, Hispanic = 78, Other = 41
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Majority severe pain at baseline (some moderate, very few mild) Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=691) Intervention 1: Strong opioids (oral) - Tapentadol. Two groups: Tapentadol ER 100-250mg given twice daily; and Oxyocodone HCl CR 20-50mg given twice daily. Duration 12 weeks. Concurrent medication/care: Additional analgesic medication was not allowed during the maintenance period (except paracetamol ≤1000mg/day, maximum, 3 consecutive days when deemed necessary for relief of pain unrelated to the index joint). Indirectness: No indirectness Comments: These two groups were reported separately but were combined for the analysis due to class effect as agreed in the protocol (n=339) Intervention 2: Placebo. Matching placebo twice daily. Duration 12 weeks. Concurrent medication/care: Additional analgesic medication was not allowed during the maintenance period (except paracetamol ≤1000mg/day, maximum, 3 consecutive days when deemed necessary for relief of pain unrelated to the index joint). Indirectness: No indirectness
Funding	Study funded by industry (Study funded by Johnson & Johnson Pharmaceutical Research & Development, L.L.C.M.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TAPENTADOL AND OXYCODONE versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D Health status index change from baseline at 12 weeks; Group 1: mean 0.16 (SD 0.37); n=686, Group 2: mean 0.1 (SD 0.37); n=337; EQ-5D 0-1 Top=High is good outcome; Comments: Reports least square means and standard error. Tapentadol and oxycodone were combined for the analysis. Reported tapentadol: 0.2 (0.02). Reported oxycodone: 0.1 (0.02). Reported placebo: 0.1 (0.02). Baseline values not reported.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, body weight, BMI, baseline pain intensity category; Group 1 Number missing: 387, Reason: Tapentadol: 346 randomised. 344 took at least 1 dose. 163 discontinued (patient choice = 50, lost

to follow up = 5, adverse event = 61, lack of efficacy = 15, study drug non-compliance = 6, other = 26). Oxycodone: 345 randomised. 342 took at least 1 dose. 224 discontinued from study (patient choice = 48, adverse event = 140, death = 1, lack of efficacy = 7, study drug non-compliance = 7, other = 21); Group 2 Number missing: 134, Reason: 339 randomised. 337 took at least 1 dose. 134 discontinued from study (patient choice = 43, lost to follow up = 3, adverse event = 22, lack of efficacy = 35, study drug non-compliance = 4, other = 27).

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -1.12 (SD 0.672); n=241, Group 2: mean -0.88 (SD 0.691); n=158; Comments: Reports least square means and standard error. Tapentadol and oxycodone were combined for the analysis. Reported tapentadol: -1.16 (0.055). Reported oxycodone: -1.05 (0.070). Reported placebo: -0.88 (0.055). Baseline values not reported.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, body weight, BMI, baseline pain intensity category; Group 1 Number missing: 387, Reason: Tapentadol: 346 randomised. 344 took at least 1 dose. 163 discontinued (patient choice = 50, lost to follow up = 5, adverse event = 61, lack of efficacy = 15, study drug non-compliance = 6, other = 26). Oxycodone: 345 randomised. 342 took at least 1 dose. 224 discontinued from study (patient choice = 48, adverse event = 140, death = 1, lack of efficacy = 7, study drug non-compliance = 7, other = 21); Group 2 Number missing: 134, Reason: 339 randomised. 337 took at least 1 dose. 134 discontinued from study (patient choice = 43, lost to follow up = 3, adverse event = 22, lack of efficacy = 35, study drug non-compliance = 4, other = 27).

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -1.04 (SD 0.67); n=241, Group 2: mean -0.83 (SD 0.69); n=158; WOMAC physical function subscale 0-4 Top=High is poor outcome; Comments: Reports least square means and standard error. Tapentadol and oxycodone were combined for the analysis. Reported tapentadol: -1.04 (0.055). Reported oxycodone: -1.04 (0.070). Reported placebo: -0.83 (0.055). Baseline values not reported.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, body weight, BMI, baseline pain intensity category; Group 1 Number missing: 387, Reason: Tapentadol: 346 randomised. 344 took at least 1 dose. 163 discontinued (patient choice = 50, lost to follow up = 5, adverse event = 61, lack of efficacy = 15, study drug non-compliance = 6, other = 26). Oxycodone: 345 randomised. 342 took at least 1 dose. 224 discontinued from study (patient choice = 48, adverse event = 140, death = 1, lack of efficacy = 7, study drug non-compliance = 7, other = 21); Group 2 Number missing: 134, Reason: 339 randomised. 337 took at least 1 dose. 134 discontinued from study (patient choice = 43, lost to follow up = 3, adverse event = 22, lack of efficacy = 35, study drug non-compliance = 4, other = 27).

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders (including constipation, nausea, vomiting, dry mouth, diarrhoea) at 12 weeks; Group 1: 378/686, Group 2: 88/337; Comments: Tapentadol: Gastrointestinal disorders total = 148, constipation = 65, nausea = 74, vomiting = 18, dry mouth = 22, diarrhoea = 16. Oxycodone: Gastrointestinal disorders total = 230, constipation = 126, nausea = 125, vomiting = 61, dry mouth = 15, diarrhoea = 17. Placebo: Gastrointestinal disorders total = 88, constipation = 22, nausea = 23, vomiting = 11, dry mouth = 8, diarrhoea = 20
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, body weight, BMI, baseline pain intensity category; Group 1 Number missing: 387, Reason: Tapentadol: 346 randomised. 344 took at least 1 dose. 163 discontinued (patient

choice = 50, lost to follow up = 5, adverse event = 61, lack of efficacy = 15, study drug non-compliance = 6, other = 26). Oxycodone: 345 randomised. 342 took at least 1 dose. 224 discontinued from study (patient choice = 48, adverse event = 140, death = 1, lack of efficacy = 7, study drug non-compliance = 7, other = 21); Group 2 Number missing: 134, Reason: 339 randomised. 337 took at least 1 dose. 134 discontinued from study (patient choice = 43, lost to follow up = 3, adverse event = 22, lack of efficacy = 35, study drug non-compliance = 4, other = 27). Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system disorders (including somnolence, dizziness, and headache) at 12 weeks; Group 1: 302/686, Group 2: 84/337; Comments: Tapentadol: Nervous system disorders = 138; somnolence = 37; dizziness = 61; headache = 51. Oxycodone: Nervous system disorders = 164; somnolence = 67; dizziness = 65; headache = 50. Placebo: Nervous system disorders = 84; somnolence = 14; dizziness = 16; headache = 56 Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, body weight, BMI, baseline pain intensity category; Group 1 Number missing: 387, Reason: Tapentadol: 346 randomised. 344 took at least 1 dose. 163 discontinued (patient choice = 50, lost to follow up = 5, adverse event = 61, lack of efficacy = 15, study drug non-compliance = 6, other = 26). Oxycodone: 345 randomised. 342 took at least 1 dose. 224 discontinued from study (patient choice = 48, adverse event = 140, death = 1, lack of efficacy = 7, study drug non-compliance = 7, other = 21); Group 2 Number missing: 134, Reason: 339 randomised. 337 took at least 1 dose. 134 discontinued from study (patient choice = 43, lost to follow up = 3, adverse event = 22, lack of efficacy = 35, study drug non-compliance = 4, other = 27).	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months

Study	Altman 2007[4]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=483)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Presence of symptomatic idiopathic osteoarthritis of the hip or knee for a minimum of 6 months with a history of hip or knee pain requiring the use of NSAIDs, paracetamol or other analgesic agent on a regular basis (3 or more days per week) for at least 3 months before the screening visit. If qualifying with OA of the knee, they had to have knee pain and radiographic osteophytes and fulfilled at least one of the following 3 criteria: morning stiffness of less than 30 min of duration, crepitus on motion, at least 40 years of age. If qualifying with OA of the hip, people must have had hip pain, radiographic femoral and/or acetabular osteophytes, and radiographic joint-space narrowing as established by the American College of Rheumatology criteria for idiopathic osteoarthritis of the hip
Stratum	Other:
Subgroup analysis within study	Not applicable
Inclusion criteria	40 years of age or older who experienced at least moderate pain when not taking any analgesic medication for osteoarthritis of the hip or knee. Presence of symptomatic idiopathic osteoarthritis of the hip or knee for a minimum of 6 months with a history of hip or knee pain requiring the use of NSAIDs, paracetamol, or other analgesic agent on a regular basis (3 or more days per week) for at least 3 months before the screening visit. People must have also had a history of positive therapeutic benefit with paracetamol use for osteoarthritis pain. In addition, people must have reported maximum osteoarthritis pain intensity experienced during the 24 hour prior to the baseline visit at a pain level of moderate or moderately severe on a 5-point Likert scale. If qualifying with osteoarthritis of the knee, they had to have knee pain and radiographic osteophytes and fulfilled at least one of the following 3 criteria: morning stiffness of less than 30 min of duration, crepitus on motion, at least 40 years of age. If qualifying with osteoarthritis of the hip, people must have had hip pain, radiographic femoral and/or acetabular osteophytes, and radiographic joint-space narrowing as established by the American College of Rheumatology criteria for idiopathic osteoarthritis of the hip. People must also have demonstrated an increase in the WOMAc pain subscale score of at least 20% relative to the screening visit score. All

	women of childbearing potential were required to have a negative urine pregnancy test and used an effective method of birth control during the study.
Exclusion criteria	History of surgery or major trauma to the study joint in the 6 months prior to the screening visit; people taking analgesic therapy for other indications, those taking anticoagulants, psychotherapeutic agents, aspirin in daily doses greater than 325mg, or statin class hypolipidaemic agents in doses that had not been stabilised within 3 months of the screening visit; people taking glucosamine, chondroitin sulfate, or shark cartilage in doses that had not been stabilised within 6 months of the screening visit; those with known alcohol abuse, intravenous drug use, drug dependency, history of psychiatric illness in the previous 12 months; administration of oral corticosteroids within 2 months of screening or intraarticular or periarticular corticosteroids or hyaluronan injections into the study joint within 6 months of screening; history of gastrointestinal or hepatic disease; clinically apparent inflammation of the study joint; medical history, physical examination or radiographic evidence suggestive of other types of osteoarthritis; collagen vascular disease; fibromyalgia
Recruitment/selection of patients	People must have had a history of positive therapeutic benefit with paracetamol use for osteoarthritis pain. The requirement of a response to paracetamol was felt to be appropriate in this setting and is analogous to the "flare" design used in most trials of antiinflammatory agents. People must also have demonstrated an increase in the WOMAc pain subscale score of at least 20% relative to the screening visit score.
Age, gender and ethnicity	Age - Mean (SD): 62.2 (10.8). Gender (M:F): 160:323. Ethnicity: Caucasian = 396, African American = 42, Other = 45
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip (majority knee)).
Extra comments	Severity: Equivalent to Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=318) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol ER 1950mg daily in three divided doses or paracetamol ER 3900mg daily in three divided doses. Duration 12 weeks. Concurrent medication/care: During the washout period, people could not taken any prescription or over-the-counter NSAID, paracetamol, aspirin, or analgesic in any form. People who had pain during the washout period were permitted to take paracetamol as rescue analgesia up to 24 hours before the baseline visit. People were guided on the appropriate use of self-administered nonpharmacologic therapies for breakthrough osteoarthritis pain, and if pain relief was inadequate,

	propoxyphene HCl (maximum dose, 390mg/day) were permitted as the only rescue analgesic medication and was to be used for no more than 3 days in any 7 day period. People were not permitted to use rescue or other prescription analgesic medications within five drug half-lives before follow-up visits for efficacy assessments. Indirectness: No indirectness (n=165) Intervention 2: Placebo. Placebo in three divided doses. Duration 12 weeks. Concurrent medication/care: During the washout period, people could not taken any prescription or over-the-counter NSAID, paracetamol, aspirin, or analgesic in any form. People who had pain during the washout period were permitted to take paracetamol as rescue analgesia up to 24 hours before the baseline visit. People were guided on the appropriate use of self-administered nonpharmacologic therapies for breakthrough osteoarthritis pain, and if pain relief was inadequate, propoxyphene HCl (maximum dose, 390mg/day) were permitted as the only rescue analgesic medication and was to be used for no more than 3 days in any 7 day period. People were not permitted to use rescue or other prescription analgesic medications within five drug half-lives before follow-up visits for efficacy assessments. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by McNeil Consumer & Specialty Pharmaceuticals, Fort Washington, PA. Editorial support was provided by Scientific Therapeutics Information, Inc, Springfield, New Jersey)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -24.7 (SD 23.7); n=318, Group 2: mean -19.6 (SD 22.5); n=165; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline paracetamol 3900mg/day: 68.9 (19.7). Baseline paracetamol 1950mg/day: 67.9 (16.5). Baseline placebo: 66.3 (19.3).
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, screening height, BMI, race and baseline values for outcomes; Group 1 Number missing: 82, Reason: Paracetamol 3900mg/d: 45 withdrew. 8 adverse events, 27 lack of efficacy, 7 protocol violations, 1 withdrew consent, 1 lost to follow up, 1 other reason. Paracetamol 1950mg/d: 37 withdrew. 10 adverse events, 16 lack of efficacy, 4 protocol violation, 6 withdrew consent, 1 lost to follow up; Group 2 Number missing: 54, Reason: 54 withdrew. 8 adverse events, 34 lack of efficacy, 7 protocol violation, 2 withdrew consent, 1 lost to follow up, 2 other reason

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -21.9 (SD 23.5); n=318, Group 2: mean -17.8 (SD 22.3); n=165;

WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline paracetamol 3900mg/day: 69.1 (18.3). Baseline paracetamol 1950mg/day: 65.9 (18.9). Baseline placebo: 65.3 (19.4).
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, screening height, BMI, race and baseline values for outcomes; Group 1 Number missing: 82, Reason: Paracetamol 3900mg/d: 45 withdrew. 8 adverse events, 27 lack of efficacy, 7 protocol violations, 1 withdrew consent, 1 lost to follow up, 1 other reason. Paracetamol 1950mg/d: 37 withdrew. 10 adverse events, 16 lack of efficacy, 4 protocol violation, 6 withdrew consent, 1 lost to follow up; Group 2 Number missing: 54, Reason: 54 withdrew. 8 adverse events, 34 lack of efficacy, 7 protocol violation, 2 withdrew consent, 1 lost to follow up, 2 other reason

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Diarrhoea at 12 weeks; Group 1: 16/318, Group 2: 4/165; Comments: Paracetamol 3900mg/d = 9. Paracetamol 1950mg/d = 7. Placebo = 4.
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, screening height, BMI, race and baseline values for outcomes; Group 1 Number missing: 82, Reason: Paracetamol 3900mg/d: 45 withdrew. 8 adverse events, 27 lack of efficacy, 7 protocol violations, 1 withdrew consent, 1 lost to follow up, 1 other reason. Paracetamol 1950mg/d: 37 withdrew. 10 adverse events, 16 lack of efficacy, 4 protocol violation, 6 withdrew consent, 1 lost to follow up; Group 2 Number missing: 54, Reason: 54 withdrew. 8 adverse events, 34 lack of efficacy, 7 protocol violation, 2 withdrew consent, 1 lost to follow up, 2 other reason

Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: AST and/or ALT levels greater than 1.5 times the upper limit of normal at 12 weeks; Group 1: 9/318, Group 2: 2/165; Comments: 3 people in the paracetamol ER 3900-mg/d group had either AST and/or ALT levels greater than 3 times the upper limit of normal. 4 people in the paracetamol 3900mg/d group, 2 people in the paracetamol 1950mg/d group and 2 people in the placebo group had minor transient increases in ALT and/or AST between 1.5 and three times the upper limit of normal.
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, screening height, BMI, race and baseline values for outcomes; Group 1 Number missing: 82, Reason: Paracetamol 3900mg/d: 45 withdrew. 8 adverse events, 27 lack of efficacy, 7 protocol violations, 1 withdrew consent, 1 lost to follow up, 1 other reason. Paracetamol 1950mg/d: 37 withdrew. 10 adverse events, 16 lack of efficacy, 4 protocol violation, 6 withdrew consent, 1 lost to follow up; Group 2 Number missing: 54, Reason: 54 withdrew. 8 adverse events, 34 lack of efficacy, 7 protocol violation, 2 withdrew consent, 1 lost to follow up, 2 other reason

Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Headache at 12 weeks; Group 1: 16/318, Group 2: 5/165; Comments: Paracetamol 3900mg/d: 9. Paracetamol 1950mg/d: 7. Placebo: 5.
Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, screening height, BMI, race and baseline values for outcomes; Group 1 Number missing: 82, Reason: Paracetamol 3900mg/d: 45 withdrew. 8 adverse events, 27 lack of efficacy, 7 protocol violations, 1 withdrew consent, 1 lost to follow up, 1 other reason. Paracetamol 1950mg/d: 37 withdrew. 10 adverse events, 16 lack of

efficacy, 4 protocol violation, 6 withdrew consent, 1 lost to follow up; Group 2 Number missing: 54, Reason: 54 withdrew. 8 adverse events, 34 lack of efficacy, 7 protocol violation, 2 withdrew consent, 1 lost to follow up, 2 other reason	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months

Study	Altman 2009[3]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=385)
Countries and setting	Conducted in Switzerland, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis in their dominant hand defined by the American College of Rheumatology criteria as nodal enlargement in ≥2 of 10 joints
Stratum	Hand
Subgroup analysis within study	Not applicable
Inclusion criteria	Osteoarthritis pain in the dominant hand for ≥12 months and use of an NSAID for ≥1 episode of pain. People meeting these criteria underwent a washout period and had to achieve post-washout criteria (see recruitment/selection of patients). Posterior-anterior radiographs had to show Kellgren-Lawrence grade 1-3 changes in symptomatic joints.
Exclusion criteria	Kellgren Lawrence grade 4 osteoarthritis; secondary osteoarthritis; other rheumatic diseases; other painful nonrheumatic diseases involving the dominant hand or arm; symptomatic osteoarthritis at additional locations besides the hand(s) requiring treatment; laboratory values indicative of rheumatoid arthritis; history of other inflammatory diseases; a diagnosis of fibromyalgia; ambidextrous participants (because evaluation of treatment outcomes required assessments in dominant versus non-dominant hands).
Recruitment/selection of patients	People meeting the inclusion criteria underwent a washout period (for at least 7 days) of previous osteoarthritis medications. Randomisation to diclofenac or placebo required people to have pain in the dominant hand during the 24 hours before the baseline visit, rated as more than or equal to 40mm on an 100mm visual analogue scale. Pain in the dominant hand had to exceed pain in the nondominant hand by more than or equal to 20mm, an people taking NSAIDs at screening had to have an increase in pain in the dominant hand of more than or equal to 15mm during washout.
Age, gender and ethnicity	Age - Mean (SD): 64.1 (10.0). Gender (M:F): 89:296. Ethnicity: (Approximate mean percentages) White = 89.1%, Black = 5.4%, Asian = 0.8%, Other = 6.3%
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hand

Extra comments	Severity: Kellgren Lawrence grade 1-3 (majority grade 3). Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=198) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac sodium gel (Voltaren® Gel, consisting of diclofenac sodium in a vehicle composed of isopropyl alcohol, propylene glycol, cocoyl caprylocaprate, mineral oil, ammonia solution, perfume cream 45/3, carbomer 980, polyoxyl 20 cetostearyl ether, and purified water) 4 times daily. 2g was judged as sufficient to cover approximately half the surface of each hand.. Duration 8 weeks. Concurrent medication/care: Rescue medication (paracetamol 500mg tablet) was allowed to a maximum dose of 4g daily during washout and throughout double-blind treatment, excluding the 36 hours before each evaluation. The same rescue medication was to be used for any other pain experienced during the trial, such as headache. Indirectness: No indirectness (n=187) Intervention 2: Placebo. Just a vehicle composed of isopropyl alcohol, propylene glycol, cocoyl caprylocaprate, mineral oil, ammonia solution, perfume cream 45/3, carbomer 980, polyoxyl 20 cetostearyl ether, and purified water) 4 times daily. 2g was judged as sufficient to cover approximately half the surface of each hand.. Duration 8 weeks. Concurrent medication/care: Rescue medication (paracetamol 500mg tablet) was allowed to a maximum dose of 4g daily during washout and throughout double-blind treatment, excluding the 36 hours before each evaluation. The same rescue medication was to be used for any other pain experienced during the trial, such as headache. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Novartis Consumer Health, Inc., Nyon, Switzerland (NCT ID: NCT00171665))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN pain index at 8 weeks; Group 1: mean 27.2 (SD 26.9); n=198, Group 2: mean 22.5 (SD 27.8); n=187; AUSCAN pain index 0-100 Top=High is poor outcome; Comments: Baseline diclofenac = 66.3 (17.9). Baseline placebo = 66.8 (16.2).
Risk of bias: All domain – High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, mean age, mean height, mean weight, mean BMI, dominant hand side, Kellgren-Lawrence grade, characteristics of the hand joints and the number of people being treated with NSAIDs; Group 1 Number missing: 25, Reason: Unable to contact person = 2, adverse events = 10, unsatisfactory therapeutic effect = 8, withdrawal of consent = 4, protocol deviation = 1; Group 2 Number missing: 26, Reason: Unable to contact person = 1, unsatisfactory therapeutic effect = 13, person withdrew consent =

6, adverse events = 4, protocol deviation = 1, administrative problems = 1 Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Hand: AUSCAN functional index at 8 weeks; Group 1: mean 26.5 (SD 27.6); n=198, Group 2: mean 19.2 (SD 28); n=187; AUSCAN functional index 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 67.9 (18.8). Baseline vehicle: 66.7 (18.4). Risk of bias: All domain – High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, mean age, mean height, mean weight, mean BMI, dominant hand side, Kellgren-Lawrence grade, characteristics of the hand joints and the number of people being treated with NSAIDs; Group 1 Number missing: 25, Reason: Unable to contact person = 2, adverse events = 10, unsatisfactory therapeutic effect = 8, withdrawal of consent = 4, protocol deviation = 1; Group 2 Number missing: 26, Reason: Unable to contact person = 1, unsatisfactory therapeutic effect = 13, person withdrew consent = 6, adverse events = 4, protocol deviation = 1, administrative problems = 1 Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Hand: Gastrointestinal adverse events at 8 weeks; Group 1: 15/198, Group 2: 7/187 Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, mean age, mean height, mean weight, mean BMI, dominant hand side, Kellgren-Lawrence grade, characteristics of the hand joints and the number of people being treated with NSAIDs; Group 1 Number missing: 25, Reason: Unable to contact person = 2, adverse events = 10, unsatisfactory therapeutic effect = 8, withdrawal of consent = 4, protocol deviation = 1; Group 2 Number missing: 26, Reason: Unable to contact person = 1, unsatisfactory therapeutic effect = 13, person withdrew consent = 6, adverse events = 4, protocol deviation = 1, administrative problems = 1 Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hand: Headache at 8 weeks; Group 1: 22/198, Group 2: 19/187 Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, mean age, mean height, mean weight, mean BMI, dominant hand side, Kellgren-Lawrence grade, characteristics of the hand joints and the number of people being treated with NSAIDs; Group 1 Number missing: 25, Reason: Unable to contact person = 2, adverse events = 10, unsatisfactory therapeutic effect = 8, withdrawal of consent = 4, protocol deviation = 1; Group 2 Number missing: 26, Reason: Unable to contact person = 1, unsatisfactory therapeutic effect = 13, person withdrew consent = 6, adverse events = 4, protocol deviation = 1, administrative problems = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Ammendolia 2021[5]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=90)
Countries and setting	Conducted in Italy; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 months (6 months total follow up)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of knee osteoarthritis according to the criteria of the American College of Rheumatology; Grade 2 according to the radiographic scale of Kellgren-Lawrence
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Age range 45-60 years; diagnosis of knee osteoarthritis according to the criteria of the American College of Rheumatology; Grade 2 according to the radiographic scale of Kellgren-Lawrence; body mass index less than 30.
Exclusion criteria	Varus or valgus malalignment; neurological and cognitive disorders; recent knee surgery, physiotherapy or intra-articular infiltrations; oral corticosteroid or NSAIDs therapy in the last month; infectious arthritis; language/cognitive deficits; unable to walk unaided.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 55 (10). Gender (M:F): Define. Ethnicity: Not stated/unclear
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren-Lawrence grade 2 Duration of symptoms (mean [SD]): 13 (11) years
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine was administered as a sachet of powder in a 1500 mg oral solution, once daily without food for 2 months.. Duration 2 months. Concurrent medication/care: People were instructed to avoid taking an analgesic, corticosteroids or other NSAIDs for the duration of the study. All people received 12 sessions, 3 per week, with diode laser applied in the perirotulum area, at the level of the articular hemirimes for a duration of 20 minutes with a wavelength of 905 nanometer, power of 4.5 Watt, dose of 70 J/cm2, pulse duration of 100 nanoseconds.. Indirectness: No indirectness (n=50) Intervention 2: Placebo. Equivalent placebo for 2 months. Duration 2 months. Concurrent medication/care: People were

	instructed to avoid taking an analgesic, corticosteroids or other NSAIDs for the duration of the study. All people received 12 sessions, 3 per week, with diode laser applied in the perirotulum area, at the level of the articular hemirimes for a duration of 20 minutes with a wavelength of 905 nanometer, power of 4.5 Watt, dose of 70 J/cm2, pulse duration of 100 nanoseconds.. Indirectness: No indirectness
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Visual analogue scale - Activities of daily living at 6 months; Group 1: mean 2 (SD 1.6); n=30, Group 2: mean 3.6 (SD 2); n=42; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline glucosamine: 7.5 (0.8). Baseline placebo: 7.5 (0.9).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of arthritis and weight; Group 1 Number missing: 10, Reason: 10 poor adherence to the trial protocol; Group 2 Number missing: 8, Reason: 8 poor adherence to the trial protocol
- Actual outcome for Knee: Visual analogue scale - Activities of daily living at 2 months; Group 1: mean 2.6 (SD 1.7); n=30, Group 2: mean 3.4 (SD 2.1); n=42; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline glucosamine: 7.5 (0.8). Baseline placebo: 7.5 (0.9).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of arthritis and weight; Group 1 Number missing: 10, Reason: 10 poor adherence to the trial protocol; Group 2 Number missing: 8, Reason: 8 poor adherence to the trial protocol

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nausea and heartburn at 6 months; Group 1: 2/40, Group 2: 0/50
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of arthritis and weight; Group 1 Number missing: 10, Reason: 10 poor adherence to the trial protocol; Group 2 Number missing: 8, Reason: 8 poor adherence to the trial protocol

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 6 months; Group 1: 2/40, Group 2: 0/50
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of arthritis and weight; Group 1 Number missing: 10, Reason: 10 poor adherence to the trial protocol; Group 2 Number missing: 8, Reason: 8 poor adherence to the trial protocol

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Amundsen 1983[6]
Study type	RCT (Patient randomised; Crossover: 0 days (between treatments, has a four day washout period before the study starts))
Number of studies (number of participants)	1 (n=52)
Countries and setting	Conducted in Norway; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 30 days (10 days for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis (knee or hip) having active synovitis with pain necessitating treatment with an NSAID
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis having acute synovitis and pain necessitating treatment with an NSAID
Exclusion criteria	People with severe hepatic or renal disease, peptic ulcers or severe infection; previous intolerance to either drug; salicylate-sensitive asthma; people who had undergone surgery within the last month
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 61 (34-78). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip, knee or multisite).
Extra comments	Severity: Not stated Duration of symptoms (mean [range]): 10 (1-38) years.
Indirectness of population	No indirectness
Interventions	(n=104) Intervention 1: NSAIDs - Naproxen. Naproxen 250mg twice a day or diclofenac 50mg twice a day. Duration 10 days (for each drug). Concurrent medication/care: Paracetamol up to 1 gram four times a day was allowed as rescue medication. Indirectness: No indirectness Comments: While diclofenac and naproxen groups were reported separately they were combined in the analysis due to class effect as stated in the protocol. While there were 52 people, for the analysis of adverse events this has been doubled to maintain the correct effect size.

	(n=52) Intervention 2: Placebo. dose/quantity, brand name, extra details. Duration 10 days. Concurrent medication/care: Paracetamol up to 1 gram four times a day was allowed as rescue medication. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN AND DICLOFENAC versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Adverse events related to the gastrointestinal tract at 10 days; Group 1: 7/104, Group 2: 6/52; Comments: No definition for outcome Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight and duration of disease for everyone, and type of osteoarthritis for each group; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Vertigo at 10 days; Group 1: 0/104, Group 2: 1/52 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight and duration of disease for everyone, and type of osteoarthritis for each group; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Andelman 1983[7]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Roentgenological and clinical criteria for the diagnosis of degenerative joint disease of the hip or knee. Had to have roentgenograms taken within the past 3 months that showed definite osteophyte formation as well as one or more of the following: lipping of marginal bone; narrowing of joint space; sharpened articular margin; damaged, thickened or eburnated subchondral bone; or bone cyst.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People selected according to roentgenological and clinical criteria. Each person included in the study was required to exhibit pain that occurred with use and was partially relieved by rest of the affected joint, as well as one or more of the following in the affected joint: limitation of movement, pain at night, tenderness when pressure was applied, or stiffness in the morning. Additional requirements were a positive response to any nonsteroidal anti-inflammatory medication used in the past, as well as a "flare" of disease within 2 weeks of withdrawal from nonsteroidal anti-inflammatory medication.
Exclusion criteria	People not meeting the inclusion criteria
Recruitment/selection of patients	Required a flare within 2 weeks of withdrawal of NSAIDs. A flare of the disease was characterised by the presence of two or more of the following: moderate or worse pain at night; one or more episodes of moderate or worse pain while bearing weightl an increase of ten minutes or more in stiffness in the morning; or a patient self-evaluation of fair or worse.
Age, gender and ethnicity	Age - Mean (range): 61.3 (44-70). Gender (M:F): 7:23. Ethnicity: White = 25, Black = 5
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).

Extra comments	Severity: Not stated Duration of symptoms (mean): 7.9 years.
Indirectness of population	No indirectness
Interventions	(n=10) Intervention 1: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: Paracetamol (650mg, up to four times daily as needed) was permitted only during the washout period. Indirectness: No indirectness (n=20) Intervention 2: NSAIDs - High-dose aspirin. Aspirin 2400-4800mg/day for 12 weeks or Etodolac 100-400mg/day for 12 weeks. Duration 12 weeks. Concurrent medication/care: Paracetamol (650mg, up to four times daily as needed) was permitted only during the washout period. Indirectness: No indirectness Comments: Aspirin and etodolac groups were combined for class effect as agreed in the protocol
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: HIGH-DOSE ASPIRIN/ETODOLAC versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastrointestinal system adverse events at 12 weeks; Group 1: 9/20, Group 2: 1/10; Comments: Including... Etodolac: Constipation = 2, diarrhoea = 1. Aspirin: Nausea = 3, GI cramps = 3, Constipation = 2, Diarrhoea= 2, Indigestion = 1, vomiting = 1, dyspepsia = 1, abdominal pain = 1. Placebo: Nausea = 1, indigestion = 1.
Risk of bias: All domain – High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, mean duration of degenerative joint disease, and mean number of days to flare; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Tachycardia at 12 weeks; Group 1: 0/20, Group 2: 1/10
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, mean duration of degenerative joint disease, and mean number of days to flare; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Central nervous system adverse events at 12 weeks; Group 1: 5/20, Group 2: 5/10; Comments: Including ... Etodolac: Headache = 1, drowsiness = 1. Aspirin: Headache = 1, drowsiness = 2, dizziness = 2, syncope (mild) = 1. Placebo: Headache = 3, dizziness = 1, crying spells = 1.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, mean duration of degenerative joint

disease, and mean number of days to flare; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Anonymous 1983[8]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=306)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 week crossover trial phase followed by a 6 week open phase
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis with a typical history and physical findings
Stratum	Other:
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis with a typical history and physical findings
Exclusion criteria	People with positive tests for rheumatoid factor; those with significant renal or hepatic disease; severe dyspepsia or active peptic ulceration; concurrent anti-arthritic therapy and pregnant or nursing mothers
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 62.9 (24-84). Gender (M:F): 78:228. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Unclear, type not stated).
Extra comments	Severity: Not stated Duration of symptoms (mean): 8.1 years.
Indirectness of population	No indirectness
Interventions	(n=306) Intervention 1: NSAIDs - Piroxicam. Piroxicam 20mg (2x10mg) once daily for 2 weeks. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=306) Intervention 2: Placebo. Placebo (2x tablets) given once daily for 2 weeks. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PIROCIXAM versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain (10cm VAS) at 2 weeks; Group 1: mean 4.1 (SD 5.5); n=289, Group 2: mean 5.6 (SD 5.5); n=289; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reports mean final values and p-value. Piroxicam = 4.1. Placebo = 5.6. P-value = <0.001. Baseline value = 6.1. Calculated SE = 0.45. Calculated SD = 5.5
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Reports gender, mean duration of disease, age, weight and height.; Group 1 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time); Group 2 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastrointestinal adverse events at 2 weeks; Group 1: 27/299, Group 2: 29/299
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Reports gender, mean duration of disease, age, weight and height.; Group 1 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time); Group 2 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time)

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Central nervous system adverse events at 2 weeks; Group 1: 12/299, Group 2: 20/299
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Reports gender, mean duration of disease, age, weight and height.; Group 1 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time); Group 2 Number missing: 17, Reason: 17 people withdrew. 7 due to lack of efficacy (3 on piroxicam and 4 on placebo at the time), and 10 due to adverse events (5 on placebo and 5 on piroxicam at the time)

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study (subsidiary papers)	Asmus 2014[9] (Asmus 2013[10])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (pooled analysis of 2 trials) (n=750)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	3rd line
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the knee in a flare state, as determined by the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged ≥40 years with diagnosed, active, symptomatic osteoarthritis of the knee in a flare state: had failed prior treatment with both prescription strength naproxen (at least 750 mg/day for 2 weeks) and ibuprofen (at least 1200 mg/day for 2 weeks) within the past 5 years due to either lack of efficacy and/or tolerability; females of childbearing potential had a negative urine pregnancy test and had to be using an adequate method of contraception; if taking chronic NSAID therapy, people were to completed a wash-out period for a minimum of 2 days; people were to have a functional capacity class of I-III; a willingness to participate for 6 weeks and ability to provide informed consent
Exclusion criteria	Inflammatory arthritis or gout/pseudo-gout with an acute flare in the past 2 years; active symptomatic acute joint trauma in the index joint within the past 3 months; previous or anticipated need for surgery on the index joint (knee arthroscopy for reasons other than arthritis was permitted as long as it was performed at least 90 days prior to screening); treatment with oral (4 weeks), intramuscular (2 months), intra-articular (3 months) or soft-tissue (2 months) injection of corticosteroids or intra-articular injections of hyaluronic acid in the index joint within 9 months of first dose of study medication; use of paracetamol within 24 hours of the baseline visit; treatment with anticoagulants, lithium, glucosamine and/or chondroitin sulfate; malignancy; treatment for oesophageal, gastric, pyloric channel or duodenal ulceration; GI or cardiovascular disease; having >1.5 times the upper limit of normal for aspartate aminotransferase, alanine aminotransferase or other clinically significant laboratory abnormalities; known sensitivity to COX-2 inhibitors or related compounds; use of study drug in the past 30 days; participation in physical therapy for the index joint and use of a mobility-assisting device <6 weeks prior to the study

Recruitment/selection of patients	People who were taking previous NSAID therapy were to complete a washout period (for at least 2 days). However, they were not selected because of characteristics after this period.
Age, gender and ethnicity	Age - Mean (SD): 59.4 (10.1). Gender (M:F): 275:493. Ethnicity: White = 486, Black = 124, Asian = 61, Other = 97
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity class I-III Duration of symptoms (mean [range]): 7.9 (0.0-64.2) years
Indirectness of population	No indirectness
Interventions	(n=385) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once daily. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=383) Intervention 2: Placebo. Placebo once daily. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (The study was funded by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale (study 1) at 6 weeks; MD; -2.3 (95%CI -3.5 to -1.5) (SE: 0.42) WOMAC pain subscale 0-20 Top=High is poor outcome, Comments: Reports least square mean, SE, 95% CI and p-value for study 1 and study 2 separately. Study 1 (celecoxib, n=186. Placebo, n=184): -2.3 (0.42). 95% CI = -3.2 to -1.5. P-value = <0.001.;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race/ethnic origin, patient's assessment of arthritis pain, WOMAC total domain, and duration of ostearthritis. Does not report the baseline values for the outcomes measured.; Group 1 Number missing: 37, Reason: Study 1: 37 discontinued (5 adverse events, 6 lack of efficacy, 19 other reasons, 7 people defaulted). Study 2: 38 discontinued (6 adverse events, 8 lack of efficacy, 21 other reasons, 3 people defaulted) ; Group 2 Number missing: 65, Reason: Study 1: 65 discontinued (8 adverse events, 21 lack of efficacy, 19 other reasons, 17 other people defaulted). Study 2: 56 discontinued (12 adverse events, 29 lack of efficacy, 10 other reasons, 5 people defaulted).
- Actual outcome for Knee: WOMAC pain subscale (study 2) at 6 weeks; MD; -0.8 (95%CI -1.5 to 0.008) (SE: 0.39) WOMAC pain subscale 0-20 Top=High is poor outcome, Comments: Reports least square mean, SE, 95% CI and p-value for study 1 and study 2 separately. Study 2 (celecoxib, n=194. Placebo, n=186): -0.8 (0.39). 95% CI = -1.5 to 0.008. P-value = 0.052.;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race/ethnic origin, patient's assessment of arthritis pain, WOMAC total domain, and duration of ostearthritis. Does not report the baseline values for the outcomes measured.; Group 1 Number missing: 37, Reason: Study 1: 37 discontinued (5 adverse events, 6 lack of efficacy, 19 other reasons, 7 people defaulted). Study 2: 38 discontinued (6 adverse events, 8 lack of efficacy, 21 other reasons, 3 people defaulted) ; Group 2 Number missing: 65, Reason: Study 1: 65 discontinued (8 adverse events, 21 lack of efficacy, 19 other reasons, 17 other people defaulted). Study 2: 56 discontinued (12 adverse events, 29 lack of efficacy, 10 other reasons, 5 people defaulted). Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function subscale (study 1) at 6 weeks; MD; -7.5 (95%CI -10.3 to -4.6) (SE: 1.4) WOMAC physical function subscale 0-68 Top=High is poor outcome, Comments: Reports least square mean, SE, 95% CI and p-value for study 1 and study 2 separately. Study 1 (celecoxib, n=186. Placebo, n=184): -7.5 (1.4). 95% CI = -10.3 to -4.6. P-value = <0.001.; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race/ethnic origin, patient's assessment of arthritis pain, WOMAC total domain, and duration of ostearthritis. Does not report the baseline values for the outcomes measured.; Group 1 Number missing: 37, Reason: Study 1: 37 discontinued (5 adverse events, 6 lack of efficacy, 19 other reasons, 7 people defaulted). Study 2: 38 discontinued (6 adverse events, 8 lack of efficacy, 21 other reasons, 3 people defaulted) ; Group 2 Number missing: 65, Reason: Study 1: 65 discontinued (8 adverse events, 21 lack of efficacy, 19 other reasons, 17 other people defaulted). Study 2: 56 discontinued (12 adverse events, 29 lack of efficacy, 10 other reasons, 5 people defaulted). - Actual outcome for Knee: WOMAC physical function subscale (study 2) at 6 weeks; MD; -2.4 (95%CI -4.9 to -0.006) (SE: 1.2) WOMAC physical function subscale 0-68 Top=High is poor outcome, Comments: Reports least square mean, SE, 95% CI and p-value for study 1 and study 2 separately. Study 2 (celecoxib, n=194. Placebo, n=186): -2.4 (1.2). 95% CI = -4.9 to -0.006. P-value = 0.049.; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race/ethnic origin, patient's assessment of arthritis pain, WOMAC total domain, and duration of ostearthritis. Does not report the baseline values for the outcomes measured.; Group 1 Number missing: 37, Reason: Study 1: 37 discontinued (5 adverse events, 6 lack of efficacy, 19 other reasons, 7 people defaulted). Study 2: 38 discontinued (6 adverse events, 8 lack of efficacy, 21 other reasons, 3 people defaulted) ; Group 2 Number missing: 65, Reason: Study 1: 65 discontinued (8 adverse events, 21 lack of efficacy, 19 other reasons, 17 other people defaulted). Study 2: 56 discontinued (12 adverse events, 29 lack of efficacy, 10 other reasons, 5 people defaulted).	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Babul 2004[11]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=246)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee meeting the American College of Rheumatology diagnostic criteria, defined by knee pain and recent radiographic evidence of osteophytes plus at least one of the following: age >50 years, morning stiffness <30 minutes in duration, and/or crepitus; have involvement of at least one knee joint that has warranted treatment with paracetamol, COX-2 inhibitors, NSAIDs, tramadol, or opioid analgesics for at least 75 of 90 days prior to the study
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of at least 18 years of age with functional class I-III primary osteoarthritis of the knee meeting the American College of Rheumatology diagnostic criteria, defined by knee pain and recent radiographic evidence of osteophytes plus at least one of the following: age >50 years, morning stiffness <30 minutes in duration, and/or crepitus; have involvement of at least one knee joint that has warranted treatment with paracetamol, COX-2 inhibitors, NSAIDs, tramadol, or opioid analgesics for at least 75 of 90 days prior to the study; and a visual analogue pain intensity score of ≥40mm in the index joint
Exclusion criteria	People with uncontrolled concomitant disease or chronic conditions that might interfere with the assessment of pain and other symptoms of osteoarthritis; other prior disease or joint replacement at the index joint; likelihood of requiring a surgical procedure of the index joint(s) during the study; inflammatory arthritis, gout, pseudogout, or Paget's disease that might interfere with the assessment of response; diagnosis of chronic pain syndrome; ACR or a clinical diagnosis of fibromyalgia; inability to discontinue paracetamol, COX-2 inhibitors, NSAIDs (other than aspirin ≤325mg once day for cardiovascular prophylaxis), corticosteroids, or other analgesics for the duration of the double-blind study; use of oral, intramuscular, intravenous, or soft tissue corticosteroids within 1 month prior to the study; use of intra-articular corticosteroids in the index knee joint within 2 months prior to the study; intra-articular viscosupplementation in the index knee joint during the past 6 months, or intra-

	articular viscosupplementation in a non-index knee in the past 3 months; weight ≤100 lbs; history of clinically significant intolerance to tramadol or a known hypersensitivity to opioid analgesics; and increased risk in terms of the precautions, warnings and contraindications noted in the tramadol prescribing indications
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 61.4 (10.1). Gender (M:F): 101:145. Ethnicity: Caucasian: 82.1%, Black: 12.1%, Hispanic: 3.2%, Other: 2.4%
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III (median class II) Duration of symptoms: 12.9 (10.5) years.
Indirectness of population	No indirectness
Interventions	(n=124) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol extended release treatment was initiated at 100mg daily and increased to 200mg daily between day 4-8, and then to 300mg or 400mg daily after the first week dependent on efficacy and adverse events. Duration 12 weeks. Concurrent medication/care: Aspirin ≤325mg/day for cardiovascular prophylaxis was permitted. Paracetamol up to 2000mg per day was permitted for reasons other than for chronic pain but for no more than 3 consecutive days and not within 24 hours of a visit.. Indirectness: No indirectness (n=122) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: Aspirin ≤325mg/day for cardiovascular prophylaxis was permitted. Paracetamol up to 2000mg per day was permitted for reasons other than for chronic pain but for no more than 3 consecutive days and not within 24 hours of a visit.. Indirectness: No indirectness
Funding	Study funded by industry (This study was managed by SCIREX Corporation, Horsham, PA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; MD; -51.1 (P value: <0.001), Comments: Reports least mean square difference and p value = -51.1 (<0.001). Calculated SE: 15.34. Calculated SD: 120.3);
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, weight, duration of

osteoarthritis, functional class, and baseline values of outcomes; Group 1 Number missing: 63, Reason: Adverse events = 33, lack of efficacy = 19, other = 11; Group 2 Number missing: 59, Reason: Adverse events = 9, lack of efficacy = 45, other = 5 Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; MD; -198.5 (P-value: 0.001) WOMAC physical function subscale 0-1700 Top=High is poor outcome, Comments: Reports least mean square difference and p value = -198.5 (<0.001). Calculated SE: 59.6. Calculated SD: 467.3. Baseline tramadol: 1158.7. Baseline placebo: 1623.1.; Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, weight, duration of osteoarthritis, functional class, and baseline values of outcomes; Group 1 Number missing: 63, Reason: Adverse events = 33, lack of efficacy = 19, other = 11; Group 2 Number missing: 59, Reason: Adverse events = 9, lack of efficacy = 45, other = 5	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Baer 2005[12]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=216)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of at least one knee. Defined by deterioration and abrasion of articular cartilage (joint space narrowing) at the joint surface of the knee (medial tibiofemoral, lateral tibiofemoral or patellofemoral) demonstrated on a radiological examination carried out within the previous 3 months
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and non-pregnant women, age 40-85 years, with primary osteoarthritis of at least one knee, and a flare of pain after withdrawal of prior therapy with either an oral NSAID or paracetamol (used as least 3 days per week during the previous month). Demonstrable radiographic features of osteoarthritis. A flare was defined as an increase in total pain subscale score of at least 2 and at least 25% with a baseline total pain score of at least 6 (out of a possible 20), and a score of at least 2 (out of a possible 4) on at least one of the 5 items of the WOMAC pain subscale
Exclusion criteria	Secondary arthritis related to systemic inflammatory arthritis (including rheumatoid arthritis, psoriatic arthritis, post-infectious arthritis and metabolic arthritis, traumatic arthritis or surgical joint replacement); corticosteroid use: a) oral corticosteorid within the previous 14 days, b) intramuscular corticosteroid within 30 days, c) intra-articular corticosteroid into the study knee within 90 days, d) intra-articular corticosteroid into any other joint within 30 days, e) topical corticosteroid at the site of application within 14 days; intraarticular viscosupplementation into the study knee in the preceding 90 days; ongoing use of prohibited medication including NSAID, other oral analgesic, muscle relaxant, or low-dose antidepressant for any chronic pain management; ongoing use of glucosamine or chondroitin (unless used continuously for 90 days prior to study entry); sensitivity to diclofenac, acetylsalicylic acid or any other NSAID, paracetamol, dimethyl sulphoxide, propylene glycol, glycerine or ethanol; clinically-active renal, hepatic, or peptic ulcer disease; history of alcohol or drug abuse; lactation; concomitant skin disease at the application site; current application for disability benefits on the basis of knee osteoarthritis; fibromyalgia; other painful or

	disabling condition affecting the knee; or participation in another investigational drug trial in the previous 30 days.
Recruitment/selection of patients	People required to have a flare after withdrawal of medication before entry into the study
Age, gender and ethnicity	Age - Mean (SD): 64.8 (11.0). Gender (M:F): 94:122. Ethnicity: White = 179, Black = 11, Oriental = 5, Other = 21
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Total X-ray score (0-27) (mean [SD]): 7.3 (5.2) Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=107) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Topical diclofenac solution 1.5% (w/w) diclofenac sodium in a patented carrier containing dimethyl sulphoxide (45.5%, w/w), propylene glycol, glycerine, ethanol and water - apply 40 drops (around 1.3mL) to the affected knee four times daily for up to 6 weeks. Duration 6 weeks. Concurrent medication/care: Paracetamol (up to four 325mg tablets per day) was permitted for residual knee or other body pain throughout the treatment period, but not during the washout period prior to baseline assessment or during the week prior to final assessment at week 6. Acetylsalicylic acid (no more than 325mg/day) was permitted for cardiovascular prophylaxis.. Indirectness: No indirectness (n=109) Intervention 2: Placebo. Topical solution of a patented carrier containing dimethyl sulphoxide (45.5%, w/w), propylene glycol, glycerine, ethanol and water - apply 40 drops (around 1.3mL) to the affected knee four times daily for up to 6 weeks. No diclofenac.. Duration 6 weeks. Concurrent medication/care: Paracetamol (up to four 325mg tablets per day) was permitted for residual knee or other body pain throughout the treatment period, but not during the washout period prior to baseline assessment or during the week prior to final assessment at week 6. Acetylsalicylic acid (no more than 325mg/day) was permitted for cardiovascular prophylaxis.. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by Dimethaid Health Care Ltd.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months	

- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -5.2 (SD 5); n=105, Group 2: mean -3.3 (SD 4.3); n=107; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline diclofenac: 13.0 (3.1). Baseline placebo: 12.7 (3.2). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, X-ray score, screening pain score, baseline values of outcomes, global assessment score, and people with bilateral osteoarthritis; Group 1 Number missing: 21, Reason: 9 adverse events, 8 lack of efficacy, 4 other; Group 2 Number missing: 39, Reason: 9 adverse events, 18 lack of efficacy, 12 other Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -13.4 (SD 16.3); n=105, Group 2: mean -6.9 (SD 13.2); n=107; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline diclofenac: 40.9 (11.9). Baseline placebo: 40.3 (11.3). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, X-ray score, screening pain score, baseline values of outcomes, global assessment score, and people with bilateral osteoarthritis; Group 1 Number missing: 21, Reason: 9 adverse events, 8 lack of efficacy, 4 other; Group 2 Number missing: 39, Reason: 9 adverse events, 18 lack of efficacy, 12 other Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Melena at 6 weeks; Group 1: 0/108, Group 2: 1/109 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, X-ray score, screening pain score, baseline values of outcomes, global assessment score, and people with bilateral osteoarthritis; Group 1 Number missing: 21, Reason: 9 adverse events, 8 lack of efficacy, 4 other; Group 2 Number missing: 39, Reason: 9 adverse events, 18 lack of efficacy, 12 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Baerwald 2010[13]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=810)
Countries and setting	Conducted in Bulgaria, Canada, Germany, Poland, Spain, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis confirmed by the presence of osteophytes in radiographs performed in no more than 12 months prior to screening
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	Male or female (using contraception, postmenopausal or not pregnant), aged at least 40 years, with a diagnosis of primary hip osteoarthritis confirmed by the presence of osteophytes in radiographs performed no more than 12 months prior to screening. People had symptoms of hip pain consistent with osteoarthritis pain during the previous 3 months, and the hip was the primary source of osteoarthritis-type pain. People had global functional status class I-III according to the American College of Rheumatology classification. People had received NSAIDs or paracetamol for osteoarthritis pain for at least 20 days during the month prior to screening. To be eligible for inclusion, people had to have experienced a flare of pain at the baseline visit (defined as a score of at least 50mm for question 1 of the WOMAC pain subscale that was increased by at least 15mm compared with the screening visit). People receiving treatment with antihypertensive agents were required to have had a stable treatment regimen during the 3 months prior to screening. Stable treatment regimens of inhaled and topical corticosteroids and low dose aspirin (no more than 325mg/day) if started before screening were allowed during the study.
Exclusion criteria	If they received oral, intramuscular, or lower-limb intraarticular corticosteroids injections within the previous 3 months; were receiving or expecting to receive treatment with phosphodiesterase V inhibitors, anticoagulants, analgesics, antiinflammatory therapies and/or nitrates or any other nitric oxide-donating drugs other than the study medication between screening and the end of the study; hepatic or renal impairment; abnormal results of a 12-lead ECG; uncontrolled diabetes mellitus; gastric or duodenal ulceration; a history of significant gastroduodenal bleeding within the previous 6 months (or any length of time in Germany); low back

	pain that may have interfered with the evaluation of OA-related pain in the hip; imminent joint replacement; current medical disease that could have interfered with the evaluation of efficacy
Recruitment/selection of patients	All analgesics apart from paracetamol was discontinued at the time of the screening visit. To be eligible for inclusion, people had to have experienced a flare of pain at the baseline visit (defined as a score of at least 50mm for question 1 of the WOMAC pain subscale that was increased by at least 15mm compared with the screening visit).
Age, gender and ethnicity	Age - Mean (SD): 63.02 (9.424). Gender (M:F): 279:531. Ethnicity: White = 786, nonwhite = 24
Further population details	1. Age: Mixed (Based on range: 41.4-90.2). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: American College of Rheumatology function class I-III (median class II). Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=156) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice daily. Duration 13 weeks. Concurrent medication/care: Paracetamol was allowed as rescue medication during the study (up to 3000mg/day, not to be used within 24 hours prior to each study visit). Other analgesic and antiinflammatory therapies were discontinued. Antihypertensives, low dose aspirin (no more than 325mg/day) and inhaled or topical corticosteroids were allowed if they had a stable dose for the past 3 months.. Indirectness: No indirectness (n=323) Intervention 2: NSAIDs - Other. Naproxcinod 750mg twice a day. Duration 13 weeks. Concurrent medication/care: Paracetamol was allowed as rescue medication during the study (up to 3000mg/day, not to be used within 24 hours prior to each study visit). Other analgesic and antiinflammatory therapies were discontinued. Antihypertensives, low dose aspirin (no more than 325mg/day) and inhaled or topical corticosteroids were allowed if they had a stable dose for the past 3 months.. Indirectness: No indirectness Comments: Naproxcinod is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=331) Intervention 3: Placebo. Placebo twice a day. Duration 13 weeks. Concurrent medication/care: Paracetamol was allowed as rescue medication during the study (up to 3000mg/day, not to be used within 24 hours prior to each study visit). Other

	analgesic and antiinflammatory therapies were discontinued. Antihypertensives, low dose aspirin (no more than 325mg/day) and inhaled or topical corticosteroids were allowed if they had a stable dose for the past 3 months.. Indirectness: No indirectness
Funding	Principal author funded by industry (Dr. Baerwald had received consulting fees, speaking fees and/or honoraria from Abbott, Bristol-Myers Squibb, Essex, Pfizer, NicOx SA, Genzymes, Wyeth, and Merck, Sharp, and Bohme (less than $10,000 each). Dr. Verdecchia has received consulting fees, speaking fees, and/or honoraria from Boehringer-Ingelheim, Novartis, Sankyo and NicOx Sa (less than $10,000 each) and owns stock in NicOx SA. Dr Duquesroix, Mrs. Frayssinet and Mr Ferreira own stock or stock options in NicOx SA.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hip: WOMAC pain subscale at 13 weeks; Group 1: mean -24.31 (SD 27.9); n=156, Group 2: mean -17.97 (SD 30.6); n=331; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Converted to SD. Reported naproxen: -24.31 (2.232). Reported placebo: -17.97 (1.681). Baseline naproxen: 63.99 (15.910). Baseline placebo: 65.48 (15.671).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, American College of Rheumatology global functional status, hypertension and baseline values for outcomes; Group 1 Number missing: 34, Reason: 34 discontinued. 5 lack of efficacy, 14 adverse events, 7 withdrew consent, 4 violation of eligibility criteria, 3 other, 1 lost to follow up.; Group 2 Number missing: 87, Reason: 87 discontinued. 38 lack of efficacy, 19 adverse events, 18 withdrew consent, 5 violation of eligibility criteria, 3 other, 4 lost to follow up

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hip: WOMAC physical function subscale at 13 weeks; Group 1: mean -21.67 (SD 27.4); n=156, Group 2: mean -13.45 (SD 29.9); n=331; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Converted to SD. Reported naproxen: -21.67 (2.191). Reported placebo: -13.45 (1.641). Baseline naproxen: 62.54 (18.224). Baseline placebo: 62.82 (17.847).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, American College of Rheumatology global functional status, hypertension and baseline values for outcomes; Group 1 Number missing: 34, Reason: 34 discontinued. 5 lack of efficacy, 14 adverse events, 7 withdrew consent, 4 violation of eligibility criteria, 3 other, 1 lost to follow up.; Group 2 Number missing: 87, Reason: 87 discontinued. 38 lack of efficacy, 19 adverse events, 18 withdrew consent, 5 violation of eligibility criteria, 3 other, 4 lost to follow up

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hip: Gastrointestinal adverse events (including dyspepsia, upper abdominal pain, diarrhoea, nausea) at 13 weeks; Group 1: 30/156, Group 2: 51/330; Comments: Reports that 19.2% of people receiving naproxen and 15.5% of those receiving placebo had gastrointestinal adverse events. This included: Naproxen: Dyspepsia = 9, upper abdominal pain = 5, diarrhoea = 6, nausea = 4. Placebo: Dyspepsia = 6, upper abdominal pain = 5, diarrhoea= 10,

nausea = 12. Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, American College of Rheumatology global functional status, hypertension and baseline values for outcomes; Group 1 Number missing: 34, Reason: 34 discontinued. 5 lack of efficacy, 14 adverse events, 7 withdrew consent, 4 violation of eligibility criteria, 3 other, 1 lost to follow up.; Group 2 Number missing: 87, Reason: 87 discontinued. 38 lack of efficacy, 19 adverse events, 18 withdrew consent, 5 violation of eligibility criteria, 3 other, 4 lost to follow up Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Hip: Cardiovascular adverse events (including serious adverse events such as worsening of hypertension, hypotension, obstructive coronary disease and death) at 13 weeks; Group 1: 6/156, Group 2: 12/330; Comments: Reported naproxen: 3.8%. Reported placebo: 3.6%. 3 people in the naproxen group had serious adverse events (worsening of hypertension, hypotension and cardiac-related death). 1 person in the placebo group had obstructive coronary disease. Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, American College of Rheumatology global functional status, hypertension and baseline values for outcomes; Group 1 Number missing: 34, Reason: 34 discontinued. 5 lack of efficacy, 14 adverse events, 7 withdrew consent, 4 violation of eligibility criteria, 3 other, 1 lost to follow up.; Group 2 Number missing: 87, Reason: 87 discontinued. 38 lack of efficacy, 19 adverse events, 18 withdrew consent, 5 violation of eligibility criteria, 3 other, 4 lost to follow up Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hip: Headache at 13 weeks; Group 1: 4/156, Group 2: 14/330 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, American College of Rheumatology global functional status, hypertension and baseline values for outcomes; Group 1 Number missing: 34, Reason: 34 discontinued. 5 lack of efficacy, 14 adverse events, 7 withdrew consent, 4 violation of eligibility criteria, 3 other, 1 lost to follow up.; Group 2 Number missing: 87, Reason: 87 discontinued. 38 lack of efficacy, 19 adverse events, 18 withdrew consent, 5 violation of eligibility criteria, 3 other, 4 lost to follow up	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Bakshi 1991[14]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=314)
Countries and setting	Conducted in United Kingdom; Setting: Primary care
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical evidence of osteoarthritis of the knees, hips or hands and who required therapy with anti-inflammatory drugs
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either sex, between the ages of 60 and 80 years, with clinical evidence of osteoarthritis of the knees, hips or hands and who required therapy with a non-steroidal anti-inflammatory drug
Exclusion criteria	History of precipitation of asthma, urticaria, rhinitis or anaphylactic reaction after intake of aspirin, diclofenac or other cyclo-oxygenase inhibitors; presence or history of peptic ulcer or gastro-intestinal bleeding; haemorrhagic diathesis; renal, hepatic or cardiac insufficiency.
Recruitment/selection of patients	Recruited from 29 general practice centers over 4 months
Age, gender and ethnicity	Age - Mean (range): 68.9 (60-80 years). Gender (M:F): 126:186. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee, hip or hand).
Extra comments	Severity: Not stated Duration of symptoms (mean): 6.4 years
Indirectness of population	No indirectness
Interventions	(n=208) Intervention 1: NSAIDs - Diclofenac. Diclofenac dispersible 50mg three times day. Duration 4 weeks. Concurrent medication/care: Analgesics (other than paracetamol), non-study systemic or topical non-steroidal anti-inflammatory drugs including over-the-counter preparations, anticoagulants, sulphonylureas and systemic/topical steroids were not permitted during the study period. Any non-pharmacological therapy, such as home exercises, was standardised. Paracetamol (up to a maximum of 4g/day) was permitted as rescue analgesic for both treatment groups.. Indirectness: No indirectness

	(n=106) Intervention 2: Placebo. Matching placebo three times a day. Duration 4 weeks. Concurrent medication/care: Analgesics (other than paracetamol), non-study systemic or topical non-steroidal anti-inflammatory drugs including over-the-counter preparations, anticoagulants, sulphonylureas and systemic/topical steroids were not permitted during the study period. Any non-pharmacological therapy, such as home exercises, was standardised. Paracetamol (up to a maximum of 4g/day) was permitted as rescue analgesic for both treatment groups.. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Gastrointestinal adverse events at 4 weeks; Group 1: 28/208, Group 2: 17/106; Comments: Including... Diclofenac: Dyspepsia/flatulence = 8, abdominal pain/discomfort = 5, Nausea = 3, vomiting = 3, diarrhoea = 2, constipation = 2, mouth ulcers = 2, anorexia = 1, heartburn = 1, gastro-enteritis = 1. Placebo: Dyspepsia/flatulence = 7, abdominal pain/discomfort = 3, nausea = 5, diarrhoea = 1, constipation = 1. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, height, weight, duration of illness and worst affected joint; Group 1 Number missing: 12, Reason: Lack of efficacy = 5, adverse effects = 4, protocol violations and other reasons = 3; Group 2 Number missing: 10, Reason: Lack of efficacy = 6, adverse effects = 3, protocol violations and other reasons = 1 Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Central nervous system adverse events at 4 weeks; Group 1: 8/208, Group 2: 5/106 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, height, weight, duration of illness and worst affected joint; Group 1 Number missing: 12, Reason: Lack of efficacy = 5, adverse effects = 4, protocol violations and other reasons = 3; Group 2 Number missing: 10, Reason: Lack of efficacy = 6, adverse effects = 3, protocol violations and other reasons = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Banerjee 2016[15]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=218)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary, symptomatic osteoarthritis of the tibio-femoral compartment of the knee joint and fulfilling the American College of Rheumatology (including radiographic evidence of osteophytes in at least one radiograph taken within the previous 6 months)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Adult men and women within 40 to 60 years of age with primary, symptomatic osteoarthritis of the tibio-femoral compartment of the knee joint and fulfilling the American College of Rheumatology criteria for diagnosis of knee osteoarthritis. In people with bilateral involvement, the more symptomatic knee was taken as the 'signal knee". Radiographic examination in those included, showed presence of osteophytes in at least one tibio-femoral compartment in a radiograph taken not more than 6 months prior to the baseline visit.
Exclusion criteria	People with a history of hypersensitivity to tapentadol or etoricoxib; secondary osteoarthritis or accompanying osteoarthritis of hip and non-osteoarthritic causes of knee pain (e.g. bursitis, fibromyalgia, osteonecrosis, ect.); people with serious or uncontrolled concomitant disease of liver, kidney, heart, thyroid, gastro-intestinal system, diabetes, HIV or malignancy; people using disease modifying osteoarthritis drugs, corticosteroids or those who had participated in any other clinical trial within the past 1 month.
Recruitment/selection of patients	Screening and recruitment of people were carried out at the Rheumatology and Orthopaedics outpatient department of a tertiary care hospital during the period of July 2012 to June 2013
Age, gender and ethnicity	Age - Mean (SD): 53.6 (6.56). Gender (M:F): 113:105. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=108) Intervention 1: Strong opioids (oral) - Tapentadol. Tapentadol 100mg twice daily orally after meals for 12 weeks. Duration 12 weeks. Concurrent medication/care: People were not allowed to use any other analgesics, corticosteroids, histamine receptor antagonists, or any other medications known to interact with or potentially alter response to the study drugs. Paracetamol 500mg tablets were used as rescue medication, only if there is severe unbearable pain in the signal knee.. Indirectness: No indirectness Comments: People had to buy their own medication, therefore were not blinded (n=110) Intervention 2: NSAIDs - Etoricoxib. Etoricoxib 30mg twice daily orally after meals. Duration 12 weeks. Concurrent medication/care: People were not allowed to use any other analgesics, corticosteroids, histamine receptor antagonists, or any other medications known to interact with or potentially alter response to the study drugs. Paracetamol 500mg tablets were used as rescue medication, only if there is severe unbearable pain in the signal knee.. Indirectness: No indirectness
Funding	Funding not stated (The article states no conflicts of interest, but nothing specifically about funding)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TAPENTADOL versus ETORICOXIB

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Visual analogue scale (pain) at 12 weeks (visit 4); Group 1: mean 25.12 (SD 3.72); n=108, Group 2: mean 26.07 (SD 4.08); n=110; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline tapentadol: 66.54 (3.98). Baseline etoricoxib: 65.87 (3.90).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI and baseline values of outcomes; Group 1 Number missing: 11, Reason: 11 people were lost to follow up; Group 2 Number missing: 15, Reason: 15 people were lost to follow up

Protocol outcome 2: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Transaminase (<2 times the upper limit of normal) at 12 weeks (visit 4); Group 1: 3/108, Group 2: 3/110
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI and baseline values of outcomes; Group 1 Number missing: 11, Reason: 11 people were lost to follow up; Group 2 Number missing: 15, Reason: 15 people were lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Baraf 2010[16]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=420)
Countries and setting	Conducted in USA; Setting: Outpatient follow up and primary care
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: According to the American College of Rheumatology criteria, people had to satisfy the following conditions at the screening visit: moderate to severe pain predominant in 1 (target) knee in the past 6 months; daily oral NSAID or paracetamol use for the past 2 weeks; target knee radiograph showing Kellgren Lawrence grade 1-3 disease in the past year; target knee pain on movement ≤80mm on a visual analogue scale
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Ambulatory men and women aged ≥35 years with osteoarthritis of 1 or both knees fulfilling the American College of Rheumatology criteria and satisfying the following criteria at the baseline visit after a washout period: target knee pain on movement score of ≥50mm, with an increase of ≥5mm from screening, WOMAC pain subscale score of ≥9 (0-20 scale), abridged WOMAC subscale pain score (questions 2 and 4, scale: 0-8) higher in the target versus contralateral knee by ≥2 points.
Exclusion criteria	Knee pain that was not caused by osteoarthritis; moderate to severe hip or back pain at baseline; rheumatoid arthritis; fibromyaglia; osteoarthritis secondary to an inflammatory joint condition, such as septic arthritis, gout, pseudogout, Paget disease, hemochromatosis, Wilson disease, heritable disorders, or collagen gene mutations; people with osteoarthritis secondary to trauma with a history of articular fracture, and/or have experienced a significant knee injury during the previous 6 months and/or had undergone major knee surgery during the last year.
Recruitment/selection of patients	Washout period after which only responders were included
Age, gender and ethnicity	Age - Mean (SD): 61.4 (10.9). Gender (M:F): 153:267. Ethnicity: White = 326, Black = 67, Other = 27
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Kellgren Lawrence grade 1-3, median = 2. Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=208) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac sodium gel 4g per knee 4 times daily for 12 weeks (to be applied to all knees where there is pain, although measurements are taken from the target knee). Duration 12 weeks. Concurrent medication/care: Paracetamol (1-2 500mg tablets, every 4 hours as needed, maximum 4g/day) were allowed as rescue medication, but was to be withheld for 48 hours before each visit. Prohibitions included corticosteroids, nonstudy analgesics (except stable doses of aspirin ≤162mg/day started ≥30 days before randomisation), topical analgesics applied to the knee, and intra- or periarticular knee injections. Indirectness: No indirectness (n=212) Intervention 2: Placebo. Vehicle only (no diclofenac, but otherwise the same), 4g applied 4 times daily. Duration 12 weeks. Concurrent medication/care: Paracetamol (1-2 500mg tablets, every 4 hours as needed, maximum 4g/day) were allowed as rescue medication, but was to be withheld for 48 hours before each visit. Prohibitions included corticosteroids, nonstudy analgesics (except stable doses of aspirin ≤162mg/day started ≥30 days before randomisation), topical analgesics applied to the knee, and intra- or periarticular knee injections. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by Novartis Consumer Health, Inc., Parsippany, NJ. Medical writing services were funded by Endo Pharmaceuticuals, Inc., Chadds Ford, PA.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean 6.8 (SD 4.5); n=207, Group 2: mean 5.4 (SD 4.5); n=212; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline diclofenac: 12.9 (3.0). Baseline placebo: 12.5 (2.7).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean 21.5 (SD 15.3); n=207, Group 2: mean 16.8 (SD 15.7); n=212; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline diclofenac: 43.2 (11.6). Baseline placebo: 42.5 (10.4).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events at 12 weeks; Group 1: 11/208, Group 2: 9/212
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cardiovascular adverse events (sinus arrhythmia) at 12 weeks; Group 1: 1/208, Group 2: 0/212
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem

Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Renal adverse events (haematuria, nephrolithiasis) at 12 weeks; Group 1: 2/208, Group 2: 0/212
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem

Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 12 weeks; Group 1: 30/208, Group 2: 28/212 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, BMI, global rating of disease, spontaneous pain, pain on movement, WOMAC subscales, Kellgren Lawrence grade and number of knees treated at baseline; Group 1 Number missing: 38, Reason: 38 discontinued: 14 adverse events, 9 unsatisfactory therapeutic effect, 10 withdrew consent, 3 lost to follow up, 1 protocol deviation, 1 administrative problem; Group 2 Number missing: 40, Reason: 40 discontinued: 3 adverse events, 20 unsatisfactory therapeutic effect, 9 withdrew consent, 7 lost to follow up, 1 administrative problem	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study (subsidiary papers)	Barthel 2009[17] (Peniston 2011[146])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=492)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of osteoarthritis in 1 or both knees according to the American College of Rheumatology criteria with a radiograph of the target knee within the previous year graded at 1, 2 or 3 on the Kellgren Lawrence disease severity scale
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Ambulatory men and women ≥35 years of age with a diagnosis of osteoarthritis in 1 or both knees with symptom onset ≥6 months before screening. Knee pain due to osteoarthritis that required oral or topical treatment with NSAIDs or paracetamol must have occurred on ≥15 days of the preceding month. A baseline pain on movement score of ≥50mm on a 100-mm visual analogue scale and a baseline WOMAC pain subscale score of ≥9 on a 20-point scale.
Exclusion criteria	Baseline pain on movement score of >20mm in the contralateral knee on a 100-mm VAS, a history of osteoarthritis pain in the contralateral knee within the previous year; a history or present evidence of secondary osteoarthritis; a history of rheumatoid arthritis or laboratory values indicative of rheumatoid arthritis with subsequent diagnosis by a physician; a history of any other chronic inflammatory joint disease or fibromyalgia.
Recruitment/selection of patients	After enrollment, eligible people underwent a 1 week washout of analgesics (or at least 5 half-lives, whichever was longer) before being randomised (there was no selection based on response after this).
Age, gender and ethnicity	Age - Mean (SD): 59.5 (10.6). Gender (M:F): 165:327. Ethnicity: Reports that 75.2% in the diclofenac group and 80.3% in the placebo group were white. No information on the remaining population.
Further population details	1. Age: Mixed (Based on range (35-92)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=254) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. 1% Diclofenac sodium gel (in a vehicle containing isopropyl alcohol, propylene glycol, cocoyl caprylocaprate, mineral oil, ammonia solution, perfuse cream 45/3, carbomer 980, polyoxyl 20 cetostearyl ether, and purified water). 4g applied 4 times daily for 12 weeks.. Duration 12 weeks. Concurrent medication/care: Rescue medication (paracetamol 500mg tablets) were supplied by the investigator to be used as an adjunct and to treat other aches and pains experience by people during the trial, such as headache. Doses of 1 or 2 tablets were permitted to a maximum of 8 tablets (4g) per day, with ≥4 hours between doses. Rescue medication was not to be used for at least 24 hours before assessments.. Indirectness: No indirectness (n=238) Intervention 2: Placebo. Placebo gel (just a vehicle containing isopropyl alcohol, propylene glycol, cocoyl caprylocaprate, mineral oil, ammonia solution, perfuse cream 45/3, carbomer 980, polyoxyl 20 cetostearyl ether, and purified water). 4g applied 4 times daily for 12 weeks.. Duration 12 weeks. Concurrent medication/care: Rescue medication (paracetamol 500mg tablets) were supplied by the investigator to be used as an adjunct and to treat other aches and pains experience by people during the trial, such as headache. Doses of 1 or 2 tablets were permitted to a maximum of 8 tablets (4g) per day, with ≥4 hours between doses. Rescue medication was not to be used for at least 24 hours before assessments.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Novartis Consumer Health, Parsippany, NJ. Funding of editorial support was provided by Endo Pharmaceuticals Inc, Chadds Ford, PA.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -5 (SD 4.3); n=253, Group 2: mean -5 (SD 4.3); n=238; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change scores and p-values. Reported diclofenac: -5.0, reported placebo: -4.0, p=0.01. Baseline diclofenac: 11.7 (2.4). Baseline placebo: 11.7 (2.5). Calculated SE = 0.39. Calculated SD = 4.3.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, BMI, and baseline values for

outcomes; Group 1 Number missing: 45, Reason: 45 discontinuations. 13 adverse events, 10 unsatisfactory therapeutic effect, 1 protocol deviation, 15 withdrew consent, 5 lost to follow up, 1 administrative problems; Group 2 Number missing: 60, Reason: 60 discontinuations. 9 adverse events, 16 unsatisfactory therapeutic effect, 5 protocol deviation, 16 withdrew consent, 12 lost to follow up, 2 administrative problems

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 12 weeks; Group 1: mean -15 (SD 13.7); n=253, Group 2: mean -10.9 (SD 13.7); n=238; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean change scores and p-values. Reported diclofenac: -15.0, reported placebo: -10.9, p=0.001. Baseline diclofenac: 38.0 (10.0). Baseline placebo: 37.9 (10.7). Calculated SE = 1.2. Calculated SD = 13.7.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, BMI, and baseline values for outcomes; Group 1 Number missing: 45, Reason: 45 discontinuations. 13 adverse events, 10 unsatisfactory therapeutic effect, 1 protocol deviation, 15 withdrew consent, 5 lost to follow up, 1 administrative problems; Group 2 Number missing: 60, Reason: 60 discontinuations. 9 adverse events, 16 unsatisfactory therapeutic effect, 5 protocol deviation, 16 withdrew consent, 12 lost to follow up, 2 administrative problems

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events at 12 weeks; Group 1: 15/254, Group 2: 12/238
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, BMI, and baseline values for outcomes; Group 1 Number missing: 45, Reason: 45 discontinuations. 13 adverse events, 10 unsatisfactory therapeutic effect, 1 protocol deviation, 15 withdrew consent, 5 lost to follow up, 1 administrative problems; Group 2 Number missing: 60, Reason: 60 discontinuations. 9 adverse events, 16 unsatisfactory therapeutic effect, 5 protocol deviation, 16 withdrew consent, 12 lost to follow up, 2 administrative problems

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cardiovascular adverse events at 12 weeks; Group 1: 4/254, Group 2: 1/238
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, BMI, and baseline values for outcomes; Group 1 Number missing: 45, Reason: 45 discontinuations. 13 adverse events, 10 unsatisfactory therapeutic effect, 1 protocol deviation, 15 withdrew consent, 5 lost to follow up, 1 administrative problems; Group 2 Number missing: 60, Reason: 60 discontinuations. 9 adverse events, 16 unsatisfactory therapeutic effect, 5 protocol deviation, 16 withdrew consent, 12 lost to follow up, 2 administrative problems

Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 12 weeks; Group 1: 35/254, Group 2: 34/238
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, BMI, and baseline values for outcomes; Group 1 Number missing: 45, Reason: 45 discontinuations. 13 adverse events, 10 unsatisfactory therapeutic effect, 1 protocol deviation, 15 withdrew consent, 5 lost to follow up, 1 administrative problems; Group 2 Number missing: 60, Reason: 60 discontinuations. 9 adverse events, 16 unsatisfactory therapeutic effect, 5 protocol deviation, 16 withdrew consent, 12 lost to follow up, 2 administrative problems

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Batlle-gualda 2007[18]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=169)
Countries and setting	Conducted in Spain; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis (people with radiological chondrocalcinosis were allowed), degrees II or III according to Kellgren-Lawrence classification, history of knee pain for at least 3 months in the last year
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People, aged 30-75 years, with primary knee osteoarthritis (people with radiological chondrocalcinosis were allowed), degrees II or III according to Kellgren-Lawrence classification, history of knee pain for at least 3 months in the last year, current knee pain ≥30mm on a visual analogue scale, and American College of Rheumatology functional classes I-III
Exclusion criteria	If they had any other kind of arthritis or connective tissue disease; knee trauma within the last three months; previous open intervention in the knee or being on a waiting list for joint replacement surgery; pregnancy or lactation; renal (creatinine >1.5mg/dL) or hepatic disease; concomitant serious medical condition or expected survival time less than 2 years; myocardial infarction or stroke in the last 4 months; history of peptic ulceration in the last two years or perforation, gastrectomy or upper GI bleeding; hypersensitivity to paracetamol or aceclofenac; use of corticosteroid in the past 3 months; concomitant use of oral anticoagulants, aspirin, corticosteroids, lithium, phenytoin, thyroxine or probenecid; females with childbearing potential not practicing adequate contraceptive measures and people enrolled in any other clinical trial within the previous 3 months or who were applying for disability for any reason
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.4 (6.8). Gender (M:F): 28:140. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: ACR functional class I-III. Kellgren Lawrence grade II-III. Duration of symptoms: 8.5 (6.5) years
Indirectness of population	No indirectness
Interventions	(n=82) Intervention 1: NSAIDs - Aceclofenac. One aceclofenac 100mg tablet and one placebo sachet in the morning, one placebo sachet at noon, one 100mg aceclofenac tablet with one placebo sachet at night. Duration 6 weeks. Concurrent medication/care: Antacid, anti-histamine-2 receptor antagonists, and proton pump inhibitors were allowed. Concurrent corticosteroid injection was not permitted. People were encouraged to keep the same level of physical activity and physical therapy.. Indirectness: No indirectness (n=86) Intervention 2: Paracetemol (oral) - Paracetemol. One placebo tablet and one paracetamol 1000mg sachet in the morning, one paracetamol sachet at noon, one placebo tablet with one paracetamol sachet at night. Duration 6 weeks. Concurrent medication/care: Antacid, anti-histamine-2 receptor antagonists, and proton pump inhibitors were allowed. Concurrent corticosteroid injection was not permitted. People were encouraged to keep the same level of physical activity and physical therapy.. Indirectness: No indirectness
Funding	Study funded by industry (This study was undertaken and supported by Almirall Prodesfarma, S.A., Barcelona, Spain)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACECLOFENAC versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain, VAS at 6 weeks; Group 1: mean -18.34 (SD 24.86); n=82, Group 2: mean -10.7 (SD 22.31); n=86; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline aceclofenac: 62.19 (20.09). Paracetamol: 62.40 (16.97).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, female, education level, BMI, duration of osteoarthritis, American College of Rheumatology functional class, radiological classification, heart rate, systolic blood pressure, diastolic blood pressure; Group 1 Number missing: 5, Reason: Lack of efficacy = 1, adverse events = 4; Group 2 Number missing: 17, Reason: Lack of efficacy = 8, adverse events = 4, personal reasons = 3, protocol violation = 1, lost to follow up = 1

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function at 6 weeks; MD; 3.14 (95%CI 0.66 to 5.61) (p-value: 0.013) WOMAC physical function subscale 0-68 Top=High is poor outcome, Comments: Does not report baseline;
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, female, education level, BMI, duration of osteoarthritis, American College of Rheumatology functional class, radiological classification, heart rate, systolic blood pressure, diastolic blood pressure; Group 1 Number missing: 5, Reason: Lack of efficacy = 1, adverse events = 4; Group 2 Number missing: 17, Reason: Lack of efficacy = 8, adverse events = 4, personal reasons = 3, protocol violation = 1, lost to follow up = 1 Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal adverse events (mild and moderate) at 6 weeks; Group 1: 22/82, Group 2: 12/86; Comments: Aceclofenac: 18 (mild), 4 (moderate). Paracetamol: 12 (mild) Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, female, education level, BMI, duration of osteoarthritis, American College of Rheumatology functional class, radiological classification, heart rate, systolic blood pressure, diastolic blood pressure; Group 1 Number missing: 5, Reason: Lack of efficacy = 1, adverse events = 4; Group 2 Number missing: 17, Reason: Lack of efficacy = 8, adverse events = 4, personal reasons = 3, protocol violation = 1, lost to follow up = 1 Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Elevated gamma glutamyl transferase at 6 weeks; Group 1: 0/82, Group 2: 1/86; Comments: Paracetamol: 1 mild elevation Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, female, education level, BMI, duration of osteoarthritis, American College of Rheumatology functional class, radiological classification, heart rate, systolic blood pressure, diastolic blood pressure; Group 1 Number missing: 5, Reason: Lack of efficacy = 1, adverse events = 4; Group 2 Number missing: 17, Reason: Lack of efficacy = 8, adverse events = 4, personal reasons = 3, protocol violation = 1, lost to follow up = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Beaulieu 2008[19]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=129)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis defined by the presence of pain of at least moderate severity, stiffness, disability and bony crepitus of the hip and/or knee joint
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and nonpregnant women between the ages of 35 and 75 years with primary osteoarthritis t requiring the use of NSAIDs, paracetamol or opioid analgesics, on a scheduled or as-needed basis, for at least 3 months before the study. As well, they needed radiographic evidence of joint degeneration, from no more than 6 months before enrolment, in the medial and/or lateral tibiofemoral compartment or in the hip, with minimum severity of grade 2. People with more advanced radiographic grades were also eligible, as long as they were not awaiting surgery.
Exclusion criteria	People with intolerance to any opioid or NSAID, or a history of drug or alcohol abuse; people with the following medical conditions: renal or hepatic impairment, secondary osteoarthritis, significant pain of alternate etiology, shortened gastrointestinal transit time, peptic ulcer disease, inflammatory disease of the gastrointestinal tract, a history of seizures or a recognised risk of seizures; people receiving corticosteroids, viscosupplementation, monoamine oxidase inhibitors, carbemazepine, quinidine, antidepressants, neuroleptics, cyclobenzaprine or promethazine.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.1 (8.7). Gender (M:F): 86:42. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).
Extra comments	Severity: Not stated explicitly. At least radiographic grade 2. Duration of symptoms (mean [SD]): 10.7 (9.4) years
Indirectness of population	No indirectness

Interventions	(n=62) Intervention 1: Strong opioids (oral) - Tramadol. Oral controlled release tramadol 200mg and a placebo diclofenac tablet uptitrated to a maximum daily dose. Duration 6 weeks. Concurrent medication/care: Breakthrough pain was managed with 325mg to 650mg paracetamol every 4 to 6 hours, as required. Indirectness: No indirectness (n=66) Intervention 2: NSAIDs - Diclofenac. Diclofenac 75mg each morning and placebo tramadol uptitrated to the maximum tolerated dose. Duration 6 weeks. Concurrent medication/care: Breakthrough pain was managed with 325mg to 650mg paracetamol every 4 to 6 hours, as required. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a research grant from Purdue Pharma, Canada)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus DICLOFENAC

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 6 weeks; Group 1: mean 73.2 (SD 99.9); n=45, Group 2: mean 80.2 (SD 108.1); n=52; WOMAC pain subscale 0-500 Top=High is poor outcome
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean duration of osteoarthritis and mean number of joint involved; Group 1 Number missing: 17, Reason: 17 withdrew. 2 protocol violations, 1 intermittent illness, 10 adverse events, 1 voluntary withdrawal, 3 inadequate pain control; Group 2 Number missing: 14, Reason: 14 withdrew. 1 intermittent illness, 10 adverse events, 1 voluntary withdrawal, 1 inadequate pain control, and 1 other

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 6 weeks; Group 1: mean 257 (SD 354.4); n=45, Group 2: mean 247.4 (SD 379.5); n=52; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Baseline tramadol: 890.8 (313.0). Baseline diclofenac: 854.5 (392.1).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean duration of osteoarthritis and mean number of joint involved; Group 1 Number missing: 17, Reason: 17 withdrew. 2 protocol violations, 1 intermittent illness, 10 adverse events, 1 voluntary withdrawal, 3 inadequate pain control; Group 2 Number missing: 14, Reason: 14 withdrew. 1 intermittent illness, 10 adverse events, 1 voluntary withdrawal, 1 inadequate pain control, and 1 other

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal

	and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Bensen 1999[20]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1003)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee fulfilling the American College of Rheumatology clinical criteria, function class of I-III. Required to have symptomatic osteoarthritis as evidence by a defined worsening of the signs and symptoms of the disease following discontinuation of treatment with NSAIDs or other analgesic medications or to have met other criteria if the person was not receiving treatment.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients after 18 year or older were eligible to participate if they fulfilled the American College of Rheumatology clinical criteria for a diagnosis of primary osteoarthritis of the knee and were in a functional class I-III. Required to have symptomatic osteoarthritis as evidence by a defined worsening of the signs and symptoms of the disease following discontinuation of treatment with NSAIDs or other analgesic medications or to have met other criteria if the person was not receiving treatment.
Exclusion criteria	People who presented with active concomitant GI tract, renal, hepatic, or coagulation disorders; malignancy (unless removed surgically with no recurrence within 5 years); oesophageal or gastroduodenal ulceration within 30 days prior to receiving the study drug (those with a history of GI tract bleeding were not excluded); people diagnosed as having any inflammatory arthritis, gout, or acute trauma of the knee; those with a known hypersensitivity to NSAIDs or sulfonamides
Recruitment/selection of patients	Required to have symptomatic osteoarthritis as evidence by a defined worsening of the signs and symptoms of the disease following discontinuation of treatment with NSAIDs or other analgesic medications or to have met other criteria if the person was not receiving treatment. After the 2-7 day washout period, a person's osteoarthritis was considered symptomatic if the person and physician's global assessment scores were "fair", "poor", or "very poor" and if 3 of the following 4 criteria were present: 1) a patient's assessment of arthritis pain measurement of 40mm or higher; 2) an increase of 2 points or more in the OA Severity Index from the screening assessment; 3) an

	increase from the screening visit of 1 grade or more in the patient's global assessment; 4) an increase from the screening visit of 1 grade or more in the physician's global assessment. For people not receiving NSAID or analgesic therapy and who had uncontrolled osteoarthritis, 3 of the following 4 conditions were necessary for randomisation at the baseline visit: 1) a person's assessment of arthritis pain measurement of 40mm or higher; 2) an osteoarthritis severity index score of 7 or more; 3) a patient's global assessment grade of poor or very poor; 4) a physician's global assessment grade of poor or very poor
Age, gender and ethnicity	Age - Mean (range): 62 (21-87). Gender (M:F): 281:722. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III Duration of symptoms (mean [SD]): 10 (8) years
Indirectness of population	No indirectness
Interventions	(n=800) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day or Celecoxib 50mg, 100mg or 200mg twice a day. Duration 12 weeks. Concurrent medication/care: Stable doses of aspirin, 325mg/d or less, and paracetamol, up to 2g/d, taken for no longer than 3 consecutive days - except during the 48-hour period prior to arthritis assessment - were allowed. The use of NSAIDs, oral or injectable corticosteroids, and anticoagulants was prohibited. Indirectness: No indirectness Comments: Celecoxib and naproxen groups were combined for class effect as agreed in the protocol (n=203) Intervention 2: Placebo. Placebo twice a day. Duration 12 weeks. Concurrent medication/care: Stable doses of aspirin, 325mg/d or less, and paracetamol, up to 2g/d, taken for no longer than 3 consecutive days - except during the 48-hour period prior to arthritis assessment - were allowed. The use of NSAIDs, oral or injectable corticosteroids, and anticoagulants was prohibited. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported in part by G.D. Searle & Co.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Average pain in the past 24 hours at 12 weeks; Group 1: mean -2.2 (SD 2.4); n=574, Group 2: mean -1.3 (SD 2.4); n=146;

American Pain Society Patient Outcome Questionnaire (VAS) 0-10 Top=High is poor outcome; Comments: Reports mean and standard error. Reported celecoxib 50mg BID = -1.7 (0.2). Reported celecoxib 100mg BID = -2.2 (0.2). Reported celecoxib 200mg BID = -2.3 (0.2). Reported naproxen = -2.4 (0.2). Reported placebo = -1.3 (0.2). Calculated SD from this. Calculated SD celecoxib 50mg BID = 2.4. Calculated SD celecoxib 100mg BID = 2.4. Calculated SD celecoxib 200mg BID = 2.4. Calculated SD naproxen = 2.4. Calculated SD placebo = 2.4.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, number of people over 70 years old, mean duration of disease, patient's global assessment and physicians global assessment, and baseline values for WOMAC composite score and osteoarthritis severity index; Group 1 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5); Group 2 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal tract adverse events at 12 weeks; Group 1: 222/800, Group 2: 45/203; Comments: Celecoxib 50mg: GI adverse events = 28%, Celecoxib 100mg: GI adverse events = 27%, Celecoxib 200mg: GI adverse events = 24%, Naproxen: GI adverse events = 32%, Placebo: GI adverse events = 22%
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, number of people over 70 years old, mean duration of disease, patient's global assessment and physicians global assessment, and baseline values for WOMAC composite score and osteoarthritis severity index; Group 1 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5); Group 2 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5)
- Actual outcome for Knee: Bleeding episode, perforation or gastric outlet obstruction at 12 weeks; Group 1: 2/800, Group 2: 0/203; Comments: Naproxen: 1 person with 2 gastric ulcers and a pyloric channel ulcer accompanied by melena and haematemesis. Celecoxib 50mg: 1 symptomatic nonbleeding duodenal ulcer
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, number of people over 70 years old, mean duration of disease, patient's global assessment and physicians global assessment, and baseline values for WOMAC composite score and osteoarthritis severity index; Group 1 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5); Group 2 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5)

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 12 weeks; Group 1: 125/800, Group 2: 45/203; Comments: Celecoxib 50mg: 19%, Celecoxib 100mg: 18%, Celecoxib 200mg: 15%, Naproxen: 14%. Placebo: 22%.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, number of people over 70 years old, mean duration of disease, patient's global assessment and physicians global assessment, and baseline values for WOMAC composite score and osteoarthritis severity index; Group 1 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5); Group 2 Number missing: -, Reason: Overall data only given. The study was completed by 569 people. Reasons for early discontinuation was treatment failure (291), adverse events (104), protocol noncompliance (33), unavailability for follow up (11), preexisting protocol violation (5)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Berry 1982[21]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (2 weeks for each active intervention and 4 weeks for placebo in total)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical evidence of osteoarthritis involving either the hip- or knee-joints
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Adult outpatients (aged 18-80 years) with clinical evidence of osteoarthritis involving either the hip or knee joints were selected
Exclusion criteria	People with significant renal, hepatic, haematological or collagen disease and people with asthma; active peptic ulceration or a recent history of severe dyspepsia
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 67.3 (51-82). Gender (M:F): 7:23. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms (mean): 7 years.
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: NSAIDs - Diclofenac. Diclofenac 50mg tablets three times daily with once daily placebo capsule or piroxicam 20mg capsule once daily with three times daily placebo tablets. Duration 4 weeks. Concurrent medication/care: People were withdrawn from all anti-inflammatory/analgesic preparations on the first day of the study period (n=30) Intervention 2: Placebo. One placebo capsule once daily and placebo tablets three times a day. Duration 4 weeks. Concurrent medication/care: People were withdrawn from all anti-inflammatory/analgesic preparations on the first day of the

	study period
Funding	Equipment / drugs provided by industry (Acknowledgement is due to Mr. Peter Mceod formerly of Pfizer Ltd. now of Warner Lambert, and to Dr. Steven Stevens Medical Director of Pfizer Ltd., for their great help and interest in designing the protocol of this trial, and to Pfizer Ltd. for provided the drugs used in the trial)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC AND PIROXICAM versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Other: Patient's assessment of pain (10cm visual analogue scale) at 4 weeks; Group 1: mean 4.04 (SD 2.88); n=21, Group 2: mean 6.25 (SD 2.77); n=22; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reports mean (SE). Reported diclofenac week 4 = 3.4 (0.73), n=11, reported diclofenac week 8 = 5.6 (0.92), n=11. Reported piroxicam week 4 = 4.8 (0.83) n=11, reported piroxicam week 8 = 2.2 (0.59) n=10. Reported placebo group A week 2 = 6.1 (0.77) n=11, reported placebo group A week 6 = 7.5 (0.62) n=11, reported placebo group B week 2 = 6.5 (0.86) n=11, reported placebo group B week 6 = 4.9 (0.87) n=11. Baseline group A = 7.3 (0.64) n=11, baseline group B = 7.0 (0.42) n=11. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height and baseline pain outcome value; Group 1 Number missing: 8, Reason: Does not report values for groups C and D, 4 people dropped out (1 for hip replacement surgery, 3 failed to keep their appointments); Group 2 Number missing: 8, Reason: Does not report values for groups C and D, 4 people dropped out (1 for hip replacement surgery, 3 failed to keep their appointments)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Berry 1983[22]
Study type	RCT (Patient randomised; Crossover: 0 days ("for ethical reasons a washout period was not included"))
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (2 weeks for each treatment)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the hip and/or knee confirmed by radiography
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Outpatients of either sex with osteoarthritis of the hip and/or knee confirmed by radiography
Exclusion criteria	People with any rheumatoid disorders; gout; a history of gastric ulcers; hepatic, renal or cardiac insufficiency; history of epilepsy; alcoholism; drug dependency; learning disability; pregnancy; or females of child-bearing age
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.3 (1.8). Gender (M:F): 11:7. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range (51-101)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee (Hip (1) or knee (17)).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=24) Intervention 1: NSAIDs - Naproxen. Naproxen 750mg per day (250mg in the morning and 500mg at night) (and a placebo dose once per day). Duration 2 weeks. Concurrent medication/care: People were asked to discontinue any analgesic or anti-inflammatory drug therapy on the day prior to commencement of the study. Throughout the study people were allowed paracetamol, up to 4g/day as a 'rescue' analgesic. No other anti-inflammatory or analgesic drug therapy was permitted. Indirectness: No indirectness (n=24) Intervention 2: Placebo. Placebo three times daily. Duration 2 weeks.

	Concurrent medication/care: People were asked to discontinue any analgesic or anti-inflammatory drug therapy on the day prior to commencement of the study. Throughout the study people were allowed paracetamol, up to 4g/day as a 'rescue' analgesic. No other anti-inflammatory or analgesic drug therapy was permitted. Indirectness: No indirectness (n=24) Intervention 3: NSAIDs - Other. Antrafenine either 150mg three times a day or 300mg three times a day. Duration 2 weeks. Concurrent medication/care: People were asked to discontinue any analgesic or anti-inflammatory drug therapy on the day prior to commencement of the study. Throughout the study people were allowed paracetamol, up to 4g/day as a 'rescue' analgesic. No other anti-inflammatory or analgesic drug therapy was permitted. Indirectness: No indirectness Comments: Antrafenine is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. However, it was reported for completeness.
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Pain on movement at 2 weeks; Group 1: mean 41 (SD 34); n=18, Group 2: mean 64 (SD 34); n=18; Visual analogue scale (pain) 0-100 Top=High is poor outcome; Comments: Reports final scores and p-values. Reported naproxen = 41.0. Reported placebo: 64.0. P-value = <0.05. Initial score = 66.5. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - High, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Crossover study. Reports sex, age, weight, height and principal joint for everyone in the study.; Group 1 Number missing: 6, Reason: 3 withdrew while taking placebo (for personal reasons unrelated to the drug therapy, 1 withdrew while taking antrafenine when she developed influenza, 1 withdrew after six weeks as they were hospitalised for hip surgery, 1 withdrew while taking naproxen due to nausea and headache; Group 2 Number missing: 6, Reason: 3 withdrew while taking placebo (for personal reasons unrelated to the drug therapy, 1 withdrew while taking antrafenine when she developed influenza, 1 withdrew after six weeks as they were hospitalised for hip surgery, 1 withdrew while taking naproxen due to nausea and headache	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-

	months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Bingham 2007[25] (Bingham 2008[24], Bingham 2011[26], Bingham 2009[27])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (pooled analysis of 2 RCTs) (n=1207)
Countries and setting	Conducted in Unknown multicentre; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (and a 12 week additional extension which undoes the randomisation and so will not be included in the analysis)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee or hip for at least 6 months who were American Rheumatology Association functional class I-III
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Otherwise healthy males or non-pregnant females, at least 40 years of age, with a diagnosis of osteoarthritis of the knee or hip for more than 6 months and were American Rheumatology Association functional class I-III. People were required to be taking an NSAID at prescription strength for at least 30 days or paracetamol 1200-4000mg strength on a regular basis (at least 25 of the last 30 days) with a history of therapeutic benefit. Eligibility required people to meet specific flare criteria upon medication washout.
Exclusion criteria	Concurrent medical or arthritis disease which could confound evaluation of efficacy (e.g. inflammatory arthritis, history of septic arthritis of the study joint, osteochondritis desiccans or osteonecrosis of the study joint, Wilson's disease, haemochromatosis, ochronosis or primary osteochondromatosis); candidates for imminent joint replacement; serum creatinine >2.0mg/dL; congestive heart failure or unstable angina; uncontrolled hypertension; stroke or transient ischaemic attack within 6 months; certain neoplastic diseases; allergy to aspirin, ibuprofen, rofecoxib, celecoxib, valdecoxib, other NSAID, paracetamol or sulpha drugs. Contraindicated prior medications within pre-specified times of initiating the study included: intravenous, intramuscular or oral corticosteroids; glucosamine and/or chondroitin sulphate; intra-articular steroids; intra-articular hyaluronans; topical, oral or systemic analgesics; warfarin, heparin and high-dose aspirin (defined as >325mg once daily); weight loss agents; appetite suppressants and chronic medications use for <1 month at a stable dose.
Recruitment/selection of patients	At screening, NSAID users had to demonstrate and assessment of pain walking on a flat surface (WOMAC question 1) of <80mm on a 100mm visual analogue scale. Paracetamol users had to demonstrate a minimum of 40mm; a score of fair, poor or

	very poor on the Investigator Global Assessment of Disease Status and a minimum of 40mm on the PGADS. Following screening, prior NSAID users discontinued treatment to allow for washout and symptom flare; paracetamol users remained on treatment. To qualify for enrollment at the flare/baseline visit, NSAID users had to demonstrate a minimum score of 40mm with an increase of 15mm on patient-assessed pain walking on a flat surface, and IGADS worsening of at least one point on a 5-point Likert scale. Paracetamol users had to demonstrate a minimum of 40mm of patient-assessed pain walking on a flat surface, fair, poor or very poor on the IGADS, and a minimum of 40mm on the PGADS.
Age, gender and ethnicity	Age - Mean (SD): 62.1 (9.6). Gender (M:F): 404:803. Ethnicity: Asian = 9, Black = 95, Hispanic = 36, White = 1059, Other = 8
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: People with multimorbidities excluded (Stated to be 'otherwise healthy'). 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American Rheumatology Association functional class I-III Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=963) Intervention 1: NSAIDs - Etoricoxib. Etoricoxib 30mg or celecoxib 200mg once a day. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (325mg or less, once daily) was allowed for cardio-protective benefit. People could continue with existing physical therapy, but were not permitted to initiate physical therapy during the study period. Indirectness: No indirectness Comments: These two groups were combined due to class effect as agreed in the protocol (n=244) Intervention 2: Placebo. Placebo once a day. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (325mg or less, once daily) was allowed for cardio-protective benefit. People could continue with existing physical therapy, but were not permitted to initiate physical therapy during the study period. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Merck and Co., Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB OR ETORICOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months

- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean 41.8 (SD 23.6); n=953, Group 2: mean 53.1 (SD 24.7); n=238; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Etoricoxib study 1: 39.6 (22.9). Etoricoxib study 2: 41.6 (23.7). Celecoxib study 1: 42.8 (22.9). Celecoxib study 2: 43.0 (24.6). Placebo study 1: 54.2 (24.6). Placebo study 2: 51.8 (24.8). Baseline etoricoxib study 1: 67.4 (16.2). Baseline etoricoxib study 2: 68.7 (16.4). Baseline celecoxib study 1: 67.5 (16.3). Baseline celecoxib study 2: 67.3 (18.7). Baseline placebo study 1: 69.1 (18.1). Baseline placebo study 2: 72.3 (17.2).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, ethnicity, prior medicine use, ARA function class, low-dose aspirin use, primary osteoarthritis joint and baseline values for outcomes; Group 1 Number missing: 166, Reason: Etoricoxib study 1: 40 withdrew. 14 lack of efficacy, 11 clinical adverse events, 9 withdrew consent, 2 protocol deviations, 1 laboratory adverse event, 3 others. Etoricoxib study 2: 32 withdrew. 15 lack of efficacy, 7 clinical adverse events, 7 withdrew consent, 1 protocol deviation, 2 other. Celecoxib study 1: 49 withdrew. 22 lack of efficacy, 11 clinical adverse events, 7 withdrew consent, 3 protocol deviations, 2 laboratory adverse events, 4 other. Celecoxib study 2: 45 withdrew. 24 lack of efficacy, 8 clinical adverse event, 5 withdrew consent, 4 protocol deviations, 1 laboratory adverse event, 4 other; Group 2 Number missing: 99, Reason: Placebo study 1: 42 withdrew. 31 lack of efficacy, 6 clinical adverse events, 3 withdrew consent, 1 protocol deviation, 1 other. Placebo study 2: 57 withdrew. 39 lack of efficacy, 12 clinical adverse events, 4 withdrew consent, 2 other.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean 43.5 (SD 23.7); n=953, Group 2: mean 54.3 (SD 24); n=237; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Etoricoxib study 1: 42.2 (22.9). Etoricoxib study 2: 44.2 (24.1). Celecoxib study 1: 44.6 (23.2). Celecoxib study 2: 43.0 (24.6). Placebo study 1: 54.6 (23.9). Placebo study 2: 53.9 (24.2). Baseline etoricoxib study 1: 65.5 (17.6). Baseline etoricoxib study 2: 67.7 (17.9). Baseline celecoxib study 1: 66.6 (17.9). Baseline celecoxib study 2: 65.8 (19.7). Baseline placebo study 1: 64.7 (18.0). Baseline placebo study 2: 65.2 (18.7).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, ethnicity, prior medicine use, ARA function class, low-dose aspirin use, primary osteoarthritis joint and baseline values for outcomes; Group 1 Number missing: 166, Reason: Etoricoxib study 1: 40 withdrew. 14 lack of efficacy, 11 clinical adverse events, 9 withdrew consent, 2 protocol deviations, 1 laboratory adverse event, 3 others. Etoricoxib study 2: 32 withdrew. 15 lack of efficacy, 7 clinical adverse events, 7 withdrew consent, 1 protocol deviation, 2 other. Celecoxib study 1: 49 withdrew. 22 lack of efficacy, 11 clinical adverse events, 7 withdrew consent, 3 protocol deviations, 2 laboratory adverse events, 4 other. Celecoxib study 2: 45 withdrew. 24 lack of efficacy, 8 clinical adverse event, 5 withdrew consent, 4 protocol deviations, 1 laboratory adverse event, 4 other; Group 2 Number missing: 99, Reason: Placebo study 1: 42 withdrew. 31 lack of efficacy, 6 clinical adverse events, 3 withdrew consent, 1 protocol deviation, 1 other. Placebo study 2: 57 withdrew. 39 lack of efficacy, 12 clinical adverse events, 4 withdrew consent, 2 other.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Discontinuation due to GI adverse events at 12 weeks; Group 1: 10/962, Group 2: 5/244; Comments: Etoricoxib study 1: 3. Etoricoxib study 2: 3. Celecoxib study 1: 2. Celecoxib study 2: 2. Placebo study 1: 0. Placebo study 2: 5.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only included people who discontinued due to adverse events; Baseline details: Reports sex, age, ethnicity, prior medicine use, ARA function class, low-dose aspirin use, primary osteoarthritis joint and baseline values for outcomes; Group 1 Number missing: 166, Reason: Etoricoxib study 1: 40 withdrew. 14 lack of efficacy, 11 clinical adverse events, 9

withdrew consent, 2 protocol deviations, 1 laboratory adverse event, 3 others. Etoricoxib study 2: 32 withdrew. 15 lack of efficacy, 7 clinical adverse events, 7 withdrew consent, 1 protocol deviation, 2 other. Celecoxib study 1: 49 withdrew. 22 lack of efficacy, 11 clinical adverse events, 7 withdrew consent, 3 protocol deviations, 2 laboratory adverse events, 4 other. Celecoxib study 2: 45 withdrew. 24 lack of efficacy, 8 clinical adverse event, 5 withdrew consent, 4 protocol deviations, 1 laboratory adverse event, 4 other; Group 2 Number missing: 99, Reason: Placebo study 1: 42 withdrew. 31 lack of efficacy, 6 clinical adverse events, 3 withdrew consent, 1 protocol deviation, 1 other. Placebo study 2: 57 withdrew. 39 lack of efficacy, 12 clinical adverse events, 4 withdrew consent, 2 other.

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Discontinuation due to oedema-related and hypertension-related adverse events and adverse events of congestive heart failure, pulmonary oedema or cardiac failure at 12 weeks; Group 1: 6/962, Group 2: 1/244; Comments: Etoricoxib study 1: Oedema-related AE: 1. Etoricoxib study 2: Oedema-related AE: 1. Celecoxib study 1: Oedema-related AE: 1. Celecoxib study 2: Oedema-related AE: 0. Placebo study 1: Oedema-related AE: 0. Placebo study 1: Oedema-related AE: 0. Etoricoxib study 1: hypertension related AE: 2. Etoricoxib study 2: hypertension related AE: 0. Celecoxib study 1: hypertension related AE: 0. Celecoxib study 2: hypertension related AE: 0. Placebo study 1: hypertension related AE: 0. Placebo study 2: hypertension related AE: 0. Etoricoxib study 1: pulmonary oedema or cardiac failure: 0. Etoricoxib study 2: pulmonary oedema or cardiac failure: 0. Celecoxib study 1: congestive heart failure, pulmonary oedema or cardiac failure: 1. Celecoxib study 2: congestive heart failure, pulmonary oedema or cardiac failure: 0. Placebo study 1: congestive heart failure, pulmonary oedema or cardiac failure: 0. Placebo study 2: congestive heart failure, pulmonary oedema or cardiac failure: 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only included people who discontinued due to adverse events; Baseline details: Reports sex, age, ethnicity, prior medicine use, ARA function class, low-dose aspirin use, primary osteoarthritis joint and baseline values for outcomes; Group 1 Number missing: 166, Reason: Etoricoxib study 1: 40 withdrew. 14 lack of efficacy, 11 clinical adverse events, 9 withdrew consent, 2 protocol deviations, 1 laboratory adverse event, 3 others. Etoricoxib study 2: 32 withdrew. 15 lack of efficacy, 7 clinical adverse events, 7 withdrew consent, 1 protocol deviation, 2 other. Celecoxib study 1: 49 withdrew. 22 lack of efficacy, 11 clinical adverse events, 7 withdrew consent, 3 protocol deviations, 2 laboratory adverse events, 4 other. Celecoxib study 2: 45 withdrew. 24 lack of efficacy, 8 clinical adverse event, 5 withdrew consent, 4 protocol deviations, 1 laboratory adverse event, 4 other; Group 2 Number missing: 99, Reason: Placebo study 1: 42 withdrew. 31 lack of efficacy, 6 clinical adverse events, 3 withdrew consent, 1 protocol deviation, 1 other. Placebo study 2: 57 withdrew. 39 lack of efficacy, 12 clinical adverse events, 4 withdrew consent, 2 other.

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	**Bhatia 2020[23]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=37)

Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with signs and symptoms of osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with signs and symptoms of osteoarthritis of the knee
Exclusion criteria	No additional information
Recruitment/selection of patients	People were recruited from the outpatient clinics of the
Age, gender and ethnicity	Age - Range: 23 to 75 years. Gender (M:F): 12:24. Ethnicity: Not stated/unclear
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated/unclear Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=24) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac gel in two forms: 1) Diclofenac lipogel, 2) Market product (Voveran(R) Emulgel(R)). People were provided with collapsible tubes (identified by a specific code number) containing the gel including 20 grams of the gel. The gel was applied to the knee twice a day for 6 weeks.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The group combines two groups: group 1: Diclofenac lipogel (N=12) and group 2: Emulgel (N=12) due to class effect as per the protocol. (n=12) Intervention 2: Placebo. Placebo gel. People were provided with collapsible tubes (identified by a specific code number) containing the gel including 20 grams of the gel. The gel was applied to the knee twice a day for 6 weeks.. Duration 6 weeks. Concurrent medication/care: No additional information.. Indirectness: No indirectness
Funding	Study funded by industry (The authors acknowledge the financial support from the M/s Life Care Innovations (P) Ltd., Gurgaon, India, for the preparation of test formulation(s). The authors are grateful to M/S Biochem Pharmaceutical Industries, Mumbai, India and M/s Lipoid GmbH, Ludwigshafen, Germany, for the ex-gratis supply of diclofenac diethylamine and phospholipids, respectively.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain at 6 weeks; Group 1: mean 8.6 (SD 3.46); n=24, Group 2: mean 13.83 (SD 4.67); n=12; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reported diclofenac lipogel: 7.46 (2.73). Reported emulgel: 9.73 (3.72). Reported placebo: 13.83 (4.67). Baseline diclofenac lipogel: 12.92 (3.5). Baseline emulgel: 13.75 (5.22). Baseline placebo: 11.67 (4.72).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not clearly report baseline values (even for outcomes, it reports the week 1 values, which means it is unclear); Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 refused to continue due to no improvement

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function at 6 weeks; Group 1: mean 36.65 (SD 10.27); n=24, Group 2: mean 53.67 (SD 13.58); n=12; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reported diclofenac lipogel: 32.85 (10.39). Reported emulgel: 40.45 (8.60). Reported placebo: 53.67 (13.58). Baseline diclofenac lipogel: 50.54 (10.60). Baseline emulgel: 51.25 (11.04). Baseline placebo: 47.25 (13.17).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not clearly report baseline values (even for outcomes, it reports the week 1 values, which means it is unclear); Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 refused to continue due to no improvement

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events at 6 weeks; Group 1: 0/24, Group 2: 0/12
Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not clearly report baseline values (even for outcomes, it reports the week 1 values, which means it is unclear); Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 refused to continue due to no improvement

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Birbara 2006[29]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (pooled analysis of 2 RCTs) (n=808)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of knee osteoarthritis for at least 6 months, with an American Rheumatology Association functional class rating of I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of at least 40 years of age with a clinical diagnosis of knee osteoarthritis for at least 6 months, and had an American Rheumatology Association functional class rating of I-III. People also had to require regular treatment with a nonselective NSAID, COX-2 selective inhibitor, or paracetamol. The eligibility criteria used for enrollment differed between people who were prior nonselective NSAID and COX-2 selective inhibitor users compared to people who were prior paracetamol users. Prior users of nonselective NSAIDs and COX-2 selective inhibitors discontinued these medications at the screening visit and returned for evaluation after an appropriate washout period of 4 to 15 days, depending on the discontinued medication. They were eligible for randomisation if they had a minimum of 40mm on 'patient assessed pain walking on a flat surface' (question 1 on the WOMAC scale) which had increased by at least 15mm since discontinuing their original osteoarthritis medication. They were also required to have a worsening in Investigator Global Assessment of Disease Status of at least 1 unit on a five-point categorical scale. People previously taking paracetamol for osteoarthritis symptoms were permitted to continue their medication until 12 hours before the study visit. Paracetamol patients were required to have a measure of at least 40mm for WOmAC pain walking on a flat surface and a rating of 'very poor', 'poor', or 'fair' on IGADS at both the screening visit and randomisation visit while on paracetamol, no flare of disease was required.
Exclusion criteria	People with a concurrent medical (or arthritic) disease that potentially could confound the evaluation of efficacy, including inflammatory arthritis and metabolic disorders associated with arthritis; people with isolated patello-femoral disease manifested by primary anterior knee pain; plans for imminent joint replacement; a history of acute injury of the study joint within the previous 2 years; clinical or laboratory evidence of

	systemic disease that would contradict the use of nonselective NSAIDs or paracetamol; serum creatinine greater than 2.0mg/dL or serum transaminases exceeding 1.5x the upper limit of normal; uncontrolled angina or congestive heart failure; myocardial infarction, cardiopulmonary bypass surgery, or angioplasty within 1 year; uncontrolled hypertension; a history of stroke or transient ischaemic attack within 1 year; a history of alcohol or substance abuse within 5 years; a history of allergic reaction to COX-2 selective inhibitors, nonselective NSAIDs, paracetamol or sulfa drugs
Recruitment/selection of patients	The eligibility criteria used for enrollment differed between people who were prior nonselective NSAID and COX-2 selective inhibitor users compared to people who were prior paracetamol users. Prior users of nonselective NSAIDs and COX-2 selective inhibitors discontinued these medications at the screening visit and returned for evaluation after an appropriate washout period of 4 to 15 days, depending on the discontinued medication. They were eligible for randomisation if they had a minimum of 40mm on 'patient assessed pain walking on a flat surface' (question 1 on the WOMAC scale) which had increased by at least 15mm since discontinuing their original osteoarthritis medication. They were also required to have a worsening in Investigator Global Assessment of Disease Status of at least 1 unit on a five-point categorical scale. People previously taking paracetamol for osteoarthritis symptoms were permitted to continue their medication until 12 hours before the study visit. Paracetamol patients were required to have a measure of at least 40mm for WOmAC pain walking on a flat surface and a rating of 'very poor', 'poor', or 'fair' on IGADS at both the screening visit and randomisation visit while on paracetamol, no flare of disease was required.
Age, gender and ethnicity	Age - Mean (SD): 60.7 (10.6). Gender (M:F): 255:553. Ethnicity: Asian = 7, Black = 49, Hispanic = 66, White = 650, Other = 30
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: American Rheumatology Association functional class rating of I-III Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=326) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day with a placebo tablet once a day. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol (maximum dose 2600mg/day) as rescue therapy for osteoarthritis pain if the study medication did not provide adequate pain control and were instructed to discontinue use 12 hours before study visits. People who used low-

	dose aspirin (81mg or less daily) for cardioprotective effects were permitted to continue low-dose aspirin use during the studies. Glucosamine and chondroitin sulfate, if taken for longer than 6 months, were also permitted if taken at the same stable dose for the duration of the studies. Use of the following agents was not permitted: intra-articular corticosteroids, sodium hyaluronate, or hylan G-F 20 (within the last 3 months); intravenous, intramuscular or oral corticosteroids (within the last month); warfarin, ticlopidine, clopidogrel, prescription weight loss agents, or appetite suppressants; and topical, oral or systemic analgesics other than the study medications or paracetamol. Indirectness: No indirectness (n=319) Intervention 2: NSAIDs - Other. Rofecoxib 12.5mg once a day and a placebo tablet once a day. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol (maximum dose 2600mg/day) as rescue therapy for osteoarthritis pain if the study medication did not provide adequate pain control and were instructed to discontinue use 12 hours before study visits. People who used low-dose aspirin (81mg or less daily) for cardioprotective effects were permitted to continue low-dose aspirin use during the studies. Glucosamine and chondroitin sulfate, if taken for longer than 6 months, were also permitted if taken at the same stable dose for the duration of the studies. Use of the following agents was not permitted: intra-articular corticosteroids, sodium hyaluronate, or hylan G-F 20 (within the last 3 months); intravenous, intramuscular or oral corticosteroids (within the last month); warfarin, ticlopidine, clopidogrel, prescription weight loss agents, or appetite suppressants; and topical, oral or systemic analgesics other than the study medications or paracetamol. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK and so was not include in the analysis as agreed in the protocol. It is reported here for completeness. (n=163) Intervention 3: Placebo. Placebo (celecoxib) and placebo (rofecoxib) once a day. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol (maximum dose 2600mg/day) as rescue therapy for osteoarthritis pain if the study medication did not provide adequate pain control and were instructed to discontinue use 12 hours before study visits. People who used low-dose aspirin (81mg or less daily) for cardioprotective effects were permitted to continue low-dose aspirin use during the studies. Glucosamine and chondroitin sulfate, if taken for longer than 6 months, were also permitted if taken at the same stable dose for the duration of the studies. Use of the following agents was not permitted: intra-articular corticosteroids, sodium hyaluronate, or hylan G-F 20 (within the last 3 months); intravenous, intramuscular or oral corticosteroids (within the last month); warfarin, ticlopidine,

	clopidogrel, prescription weight loss agents, or appetite suppressants; and topical, oral or systemic analgesics other than the study medications or paracetamol. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Merck & Co. Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -27.9 (SD 23.6); n=320, Group 2: mean -20.1 (SD 23.6); n=159; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and p-values. Calculated SD from this. Reported study 1 celecoxib: -27.30. Reported study 2 celecoxib: -28.51. Reported study 1 placebo: -19.24. Reported study 2 placebo: -20.95. Reported study 1 p-value: 0.021. Reported study 2 p-value: 0.013. Calculated standard error group 1: 3.47. Calculated SD group 1: 25.0. Calculated standard error group 2: 3.02. Calculated SD group 2: 22.1. Baseline study 1 celecoxib: 68.3 (17.0). Baseline study 2 celecoxib: 66.8 (17.3). Baseline study 1 placebo: 68.5 (17.0). Baseline study 2 placebo: 68.7 (17.6).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, medication used and baseline values of outcomes; Group 1 Number missing: 40, Reason: Study 1 celecoxib: 14 discontinued. 3 for clinical adverse experience, 2 lost to follow up, 3 withdrew consent, 5 lack of efficacy, 1 moved. Study 2: 26 discontinued. 6 for clinical adverse experience, 4 lost to follow up, 2 withdrew consent, 11 lack of efficacy, 3 other.; Group 2 Number missing: 47, Reason: Study 1 placebo: 21 discontinued. 2 clinical adverse experience, 1 withdrew consent, 17 lack of efficacy, 1 other. Study 2 placebo: 26 discontinued. 3 clinical adverse experience, 3 lost to follow up, 6 withdrew consent, 13 lack of efficacy, 1 other.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -25.1 (SD 24.3); n=321, Group 2: mean -17.2 (SD 24.2); n=159; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and p-values. Calculated SD from this. Reported study 1 celecoxib: -26.52. Reported study 2 celecoxib: -23.71. Reported study 1 placebo: -16.90. Reported study 2 placebo: -17.45. Reported study 1 p-value: 0.010. Reported study 2 p-value: 0.034. Calculated standard error group 1: 3.70. Calculated SD group 1: 26.7. Calculated standard error group 2: 2.94. Calculated SD group 2: 21.6. Baseline study 1 celecoxib: 67.5 (17.5). Baseline study 2 celecoxib: 63.1 (19.8). Baseline study 1 placebo: 65.0 (19.4). Baseline study 2 placebo: 65.1 (19.4).

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, medication used and baseline values of outcomes; Group 1 Number missing: 40, Reason: Study 1 celecoxib: 14 discontinued. 3 for clinical adverse experience, 2 lost to follow up, 3 withdrew consent, 5 lack of efficacy, 1 moved. Study 2: 26 discontinued. 6 for clinical adverse experience, 4 lost to follow up, 2 withdrew consent, 11 lack of efficacy, 3 other.; Group 2 Number missing: 47, Reason: Study 1 placebo: 21 discontinued. 2 clinical adverse experience, 1 withdrew consent, 17 lack of efficacy, 1 other. Study 2 placebo: 26 discontinued. 3 clinical adverse experience, 3 lost to follow up, 6 withdrew consent, 13 lack of efficacy, 1 other.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Knee: Gastrointestinal adverse experiences at 6 weeks; Group 1: 23/326, Group 2: 11/162 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, medication used and baseline values of outcomes; Group 1 Number missing: 40, Reason: Study 1 celecoxib: 14 discontinued. 3 for clinical adverse experience, 2 lost to follow up, 3 withdrew consent, 5 lack of efficacy, 1 moved. Study 2: 26 discontinued. 6 for clinical adverse experience, 4 lost to follow up, 2 withdrew consent, 11 lack of efficacy, 3 other.; Group 2 Number missing: 47, Reason: Study 1 placebo: 21 discontinued. 2 clinical adverse experience, 1 withdrew consent, 17 lack of efficacy, 1 other. Study 2 placebo: 26 discontinued. 3 clinical adverse experience, 3 lost to follow up, 6 withdrew consent, 13 lack of efficacy, 1 other. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Oedema-related adverse experiences causing discontinuation and Hypertension-related adverse experiences causing discontinuation at 6 weeks; Group 1: 1/326, Group 2: 2/162; Comments: Celecoxib: hypertension-related events = 1. Placebo: hypertension-related events = 1, oedema-related events = 1. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Adverse events are recorded only when they lead to discontinuation; Baseline details: Reports sex, age, race, medication used and baseline values of outcomes; Group 1 Number missing: 40, Reason: Study 1 celecoxib: 14 discontinued. 3 for clinical adverse experience, 2 lost to follow up, 3 withdrew consent, 5 lack of efficacy, 1 moved. Study 2: 26 discontinued. 6 for clinical adverse experience, 4 lost to follow up, 2 withdrew consent, 11 lack of efficacy, 3 other.; Group 2 Number missing: 47, Reason: Study 1 placebo: 21 discontinued. 2 clinical adverse experience, 1 withdrew consent, 17 lack of efficacy, 1 other. Study 2 placebo: 26 discontinued. 3 clinical adverse experience, 3 lost to follow up, 6 withdrew consent, 13 lack of efficacy, 1 other.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Bocanegra 1998[31]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=572)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee and/or hip with a functional capacity classification of 1-3
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults with symptomatic osteoarthritis of the knee and/or hip; a Functional Capacity Classification of I-III; and a documented history but not current presence of significant upper GI mucosal damage (i.e. gastric, pyloric channel, or duodenal ulcer, or more than 10 endoscopically confirmed erosions in the stomach or duodenum)
Exclusion criteria	Arthritis other than osteoarthritis; malignancy of any type; renal, hepatic, or coagulation disorder that, in the opinion of the investigator, may pose a safety risk for the patient; presence of gastric or duodenal surgery other than a simple oversew; recent use of corticosteroids (including intraarticular injections) or anticoagulants; or use of any NSAID except aspirin ≤325mg/day) or any analgesics within 3 days prior to baseline arthritis assessments.
Recruitment/selection of patients	After the screening visit, those taking NSAID or analgesics discontinued their current arthritis medication and returned to the clinic, between 3 and 14 days after discontinuation for the baseline evaluation. To be eligible for the study, at the baseline visit, people had to meet all the inclusion and none of the exclusion criteria and have worsening of osteoarthritis symptoms compared to screening evaluation, and not have an ulcer or more than 10 erosions in either the stomach or the duodenum on baseline upper GI endoscopy. Worsening of symptoms was defined as at least 2 of the following 3: 1) an increase of one grade or more since screening or a score of "poor " or "very poor" on the Physician's global assessment; 2) an increase of one grade or more since screening or a score of "poor " or "very poor" on the Patient's global assessment; 3) An increase of at least 2 points since screening, or a score of 7 or higher, on the Osteoarthritis Severity Index
Age, gender and ethnicity	Age - Mean (SD): 62.5 (10.4). Gender (M:F): 180:392. Ethnicity: Not stated

Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).
Extra comments	Severity: Functional capacity classification 1-3 Duration of symptom (mean [SD]: 11.2 (8.7) years.
Indirectness of population	No indirectness
Interventions	(n=327) Intervention 1: NSAID and gastroprotection - NSAID and misoprostol . Either diclofenac 50mg with misoprostol 200 micrograms three times a day, or diclofenac 75mg with misoprostol 200 micrograms two times a day . Duration 6 weeks. Concurrent medication/care: No additional information - all antiinflammatory drugs were not permitted (with the exception of aspirin ≤325mg/day).. Indirectness: No indirectness Comments: These two formulations of diclofenac and misoprostol were reported separately but were combined for the analysis as per the protocol (n=154) Intervention 2: NSAIDs - Diclofenac. Diclofenac 75mg twice a day. Duration 6 weeks. Concurrent medication/care: No additional information - all antiinflammatory drugs were not permitted (with the exception of aspirin ≤325mg/day).. Indirectness: No indirectness (n=91) Intervention 3: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: No additional information - all antiinflammatory drugs were not permitted (with the exception of aspirin ≤325mg/day).. Indirectness: No indirectness
Funding	Study funded by industry (Supported by G.D. Searle & Co., Skokie, Illinois, USA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC AND MISOPROSTOL versus DICLOFENAC

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Patient's assessment of arthritis pain, VAS at 6 weeks; Group 1: mean -2.89 (SD 2.99); n=327, Group 2: mean -2.87 (SD 3.08); n=154; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reported D50/M200 = -2.39 (2.77). Reported D75/M200 = -2.95 (2.80). Baseline D50/M200 6.3 (1.91). Baseline D75/M200 = 6.6 (2.03). Baseline Diclofenac = 6.6 (2.14).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulcers at 6 weeks; Group 1: 22/327, Group 2: 24/154; Comments: D50/M200: gastroduodenal ulcer = 11. D75/M200: gastroduodenal ulcer = 11. Diclofenac: gastroduodenal ulcer = 24.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Increases in AST or ALT levels to above 3 times the upper limit of normal at 6 weeks; Group 1: 5/327, Group 2: 2/154
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC AND MISOPROSTOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Patient's assessment of arthritis pain, VAS at 6 weeks; Group 1: mean -2.89 (SD 2.99); n=327, Group 2: mean -1.3 (SD 3.04); n=91; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reported D50/M200 = -2.39 (2.77). Reported D75/M200 = -2.95 (2.80). Baseline D50/M200 6.3 (1.91). Baseline D75/M200 = 6.6 (2.03). Baseline placebo = 6.4 (1.92).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulcers at 6 weeks; Group 1: 22/327, Group 2: 3/91; Comments: D50/M200: gastroduodenal ulcer = 11. D75/M200: gastroduodenal ulcer = 11. Placebo: gastroduodenal ulcer = 3.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Increases in AST or ALT levels to above 3 times the upper limit of normal at 6 weeks; Group 1: 5/327, Group 2: 0/91
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 54, Reason: 54 withdrew. 37 withdrew due to adverse events, 6 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Patient's assessment of arthritis pain, VAS at 6 weeks; Group 1: mean -2.87 (SD 3.08); n=154, Group 2: mean -1.3 (SD 3.04); n=91; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline Diclofenac = 6.6 (2.14). Baseline placebo = 6.4 (1.92).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulcers at 6 weeks; Group 1: 24/154, Group 2: 3/91; Comments: Diclofenac: gastroduodenal ulcer = 24. Placebo: gastroduodenal ulcer = 3.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Increases in AST or ALT levels to above 3 times the upper limit of normal at 6 weeks; Group 1: 2/154, Group 2: 0/91
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, affected joint, disease duration and baseline values of outcomes; Group 1 Number missing: 28, Reason: 28 withdrew. 20 withdrew due to adverse events, 3 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution); Group 2 Number missing: 21, Reason: 21 withdrew. 6 withdrew due to adverse events, 14 withdrew due to lack of efficacy, the remainder withdrew for 'miscellaneous' reasons (not given with group distribution)

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Bolten 1992[32]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=361)
Countries and setting	Conducted in Australia, Belgium, Canada, France, Germany, Greece, Luxembourg, Mexico, Portugal, United Kingdom, Venezuela; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A confirmed diagnosis of osteoarthritis of the hip and/or knee (at least clinical assessment)
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People of any sex, of the legal age of consent with a confirmed diagnosis of osteoarthritis of the hip and/or knee of at least 3 months duration, to require continuous NSAID therapy for the duration of the study and to have a functional capacity classification of I-III. Female candidates of childbearing potential were required to have a negative pregnancy test prior to starting the study and to use adequate contraception during the study.
Exclusion criteria	Arthritis other than osteoarthritis; any other rheumatic disease, psoriasis, acute joint trauma at the osteoarthritis site; any musculoskeletal disorder of the lumbosacral area, syphilitic neuropathy, ochronosis, or metabolic bone disease; candidates with significant upper gastrointestinal mucosal damage (i.e. >10 erosions in the stomach; >10 erosions in the duodenum; oesophageal, gastric, pyloric channel or duodenal ulcer); any active gastrointestinal disease; renal or hepatic disorders; or malignancy; use of corticosteroids or disease-modifying antirheumatic drugs in the past 30 days; chronic use of analgesics or NSAIDs (including aspirin).
Recruitment/selection of patients	Multicenter
Age, gender and ethnicity	Age - Mean (SD): 60.3 (12.0). Gender (M:F): 97:264. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range, 31-91). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (52), knee (208) or both (101)).
Extra comments	Severity: Functional capacity class I-IV, mean osteoarthritis severity index 11.2 (3.7) Duration of symptoms: Median 1-10 years.
Indirectness of population	No indirectness

Interventions	(n=178) Intervention 1: NSAID and gastroprotection - NSAID and misoprostol . Diclofenac with misoprostol - 50mg diclofenac sodium with 200micrograms misoprostol twice or three times a day (at the discretion of the investigator for adequate control of the person's arthritis). Duration 4 weeks. Concurrent medication/care: The use of any other NSAIDs, analgesics or antiulcer drugs during the study was prohibited. Indirectness: No indirectness (n=183) Intervention 2: NSAIDs - Diclofenac. 50mg diclofenac sodium two to three times per day (at the discretion of the investigator for adequate control of the person's arthritis) with matching placebo (to correspond with misoprotsol). Duration 4 weeks. Concurrent medication/care: The use of any other NSAIDs, analgesics or antiulcer drugs during the study was prohibited. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NSAID AND MISOPROSTOL versus DICLOFENAC

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Ulcer of any size, erosions and petechiae at 4 weeks; Group 1: 50/162, Group 2: 72/167; Comments: Diclofenac/misoprostol: Ulcer of any size = 0, >10 (?>25) erosions = 2, 11-25 erosions = 1, 6-10 erosions = 4, 1-5 erosions = 20, >10 petechiae = 2, 1-10 petechiae = 21. Diclofenac/placebo: Ulcer of any size = 6, >10 (?>25) erosions = 0, 11-25 erosions = 2, 6-10 erosions = 5, 1-5 erosions = 37, >10 petechiae = 1, 1-10 petechiae = 21.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, major joint affected, duration of arthritis, functional capacity, and baseline values for outcomes. Does not mention number of people who had twice daily or three times daily medication; Group 1 Number missing: 26, Reason: Endoscopy completed in 162 people. 3 withdrew due to protocol violations, 10 due to adverse events, 6 were lost to follow up.; Group 2 Number missing: 27, Reason: Endoscopy completed in 167 people. 3 were withdrawn due to protocol violations, 11 due to adverse events, 3 were lost to follow up

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Headache at 4 weeks; Group 1: 12/178, Group 2: 20/183
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, major joint affected, duration of arthritis, functional capacity, and baseline values for outcomes. Does not mention number of people who had twice daily or three times daily medication; Group 1 Number missing: 26, Reason: Endoscopy completed in 162 people. 3 withdrew due to protocol violations, 10 due to adverse events, 6 were lost to follow up.; Group 2 Number missing: 27, Reason: Endoscopy completed in 167 people. 3 were withdrawn due to protocol violations, 11 due to adverse events, 3 were lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Bookman 2004[33]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=248)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis in at least 1 knee, verified radiologically within the previous 6 months and at least moderate pain during the 2 weeks before random assignment to treatment.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Primary osteoarthritis in at least 1 knee, verified radiologically within the previous 6 months and scored (as normal, minimal, moderate or marked) for joint-space narrowing and marginal osteophytes in the medial, lateral and patellofemoral compartments, at least moderate pain during the 2 weeks before random assignment to treatment, as identified with the WOMAC pain subscale; age 18 to 80 years; if female could not become pregnant.
Exclusion criteria	Secondary arthritis related to syphilitic neuropathy; ochronosis; metabolic bone disease or acute trauma; sensitivity to diclofenac, acetylsalicylic acid or any other NSAID, DMSO, propylene glycol, glycerin or ethanol; clinically active renal, hepatic or peptic ulcer disease; history of alcohol or drug abuse; lactation; concomitant skin disease at the application site; corticosteroid use; use of another topical product at the application site; oral use of an analgesic or glucosamine.
Recruitment/selection of patients	People using NSAIDs or other prohibited medications underwent a 1-week washout of that medication before baseline assessment and random allocation to treatment (this was a nonflare study: an increase in pain after washout of previous medication was not required).
Age, gender and ethnicity	Age - Mean (SD): 61.8 (11.5). Gender (M:F): 91:157. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range: 18-80 years). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated

Indirectness of population	No indirectness
Interventions	(n=84) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. 1.5% wt/wt diclofenac sodium in a carrier containing DMSO (45.5% wt/wt), propylene glycol, glycerin, ethanol and water. Duration 4 weeks. Concurrent medication/care: The use of acetylsalicylic acid (no more than 325mg/day) was permitted for cardiovascular prophylaxis. Indirectness: No indirectness (n=164) Intervention 2: Placebo. Placebo gels - a vehicle-control solution consisting of the complete carrier, including DMSO (45.5% wt/wt), but no diclofenac; and a placebo solution consisting of a modified carrier with a token amount of DMSO (4.55% wt/wt) for blinding purposes but no dliclofenac. Duration 4 weeks. Concurrent medication/care: The use of acetylsalicylic acid (no more than 325mg/day) was permitted for cardiovascular prophylaxis. Indirectness: No indirectness Comments: These two groups were pooled together as they were both deemed to be placebo agents
Funding	Other author(s) funded by industry (Kate Williams and Zev Shainhouse were employees of Dimethaid Health Care Ltd when the study was conducted)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 4 weeks; Group 1: mean -3.9 (SD 4.4); n=84, Group 2: mean -2.5 (SD 3.7); n=163; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Calculated SD from this. Reported diclofenac: -3.9 (-4.8 to -2.9). Reported vehicle control: -2.5 (-3.3 to -1.7). Reported placebo: -2.5 (-3.3 to -1.7). Calculated SD diclofenac: 4.4. Calculated SD vehicle control: 3.6. Calculated SD placebo: 3.7. Baseline diclofenac: 9.1 (3.5). Baseline vehicle control: 9.3 (3.5). Baseline placebo: 9.4 (3.6).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, height, people with bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 10, Reason: Adverse events related to study solution = 5, lack of effect = 2, other medical or personal reasons = 3.; Group 2 Number missing: 29, Reason: Vehicle-control solution: Adverse events related to study solution = 3, lack of effect = 8, other medical or personal reason = 3. Placebo solution: Adverse event related to study solution = 0, lack of effect = 10, other medical or personal reason = 5.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 4 weeks; Group 1: mean -11.6 (SD 14.7); n=84, Group 2: mean -6.4 (SD 11.6); n=163; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Calculated SD from this. Reported diclofenac: -11.6 (-14.7 to -8.4). Reported vehicle control: -5.7 (-8.3 to -3.2). Reported placebo: -7.1 (-9.3 to -4.4). Calculated SD diclofenac:

14.7. Calculated SD vehicle control: 11.6. Calculated SD placebo: 11.5. Baseline diclofenac: 29.5 (13.7). Baseline vehicle control: 30.5 (11.9). Baseline placebo: 30.9 (13.1). Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, height, people with bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 10, Reason: Adverse events related to study solution = 5, lack of effect = 2, other medical or personal reasons = 3.; Group 2 Number missing: 29, Reason: Vehicle-control solution: Adverse events related to study solution = 3, lack of effect = 8, other medical or personal reason = 3. Placebo solution: Adverse event related to study solution = 0, lack of effect = 10, other medical or personal reason = 5. Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: ALT and AST elevated values at 4 weeks; Group 1: 2/82, Group 2: 5/149; Comments: Diclofenac: ALT and AST: 1, AST only: 1. Vehicle-control: ALT and AST: 2, ALT only: 2, Placebo: AST only: 1. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, height, people with bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 5, Reason: Lack of effect = 2, other medical or personal reasons = 3.; Group 2 Number missing: 26, Reason: Vehicle-control solution: Lack of effect = 8, other medical or personal reason = 3. Placebo solution: Lack of effect = 10, other medical or personal reason = 5.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Bradley 1991[35] (Bradley 2001[36])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=184)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee pain and evidence of grade 2-3 osteoarthritis on standing anteroposterior and lateral radiographs of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	All were at least 30 years old and had knee pain and evidence of grade 2 or 3 moderate osteoarthritis on standing anteroposterior and lateral radiographs of the knee. All were able to walk without assistance or assistive devices.
Exclusion criteria	People with fibromyalgia, bursitis, tendinitis, or severe neurologic or vascular disease affecting the lower extremities were excluded. None of the participants had a history of trauma, surgery, or injection of the knee with corticosteroids in the previous 3 months, an underlying inflammatory arthropathy (e.g. rheumatoid arthritis, gout, or pseudogout), or a medical condition that would have contraindicated the use of the study medications.
Recruitment/selection of patients	People were recruited from the general medicine, rheumatology and orthopaedic surgery clinics of the Indiana University School of Medicine and from the surrounding community. There was a washout period of previous medications for 3-7 days. However, all participants developed knee pain and were included in the study.
Age, gender and ethnicity	Age - Mean (SD): 56.5 (12.3). Gender (M:F): 47:137. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Radiographic grade 2-3 Duration of symptoms (mean [SD]): 9.2 (9.4) years
Indirectness of population	No indirectness
Interventions	(n=123) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 1200mg per day or 2400mg per day given as ibuprofen 150mg or 300mg tablets 8 tablets (2 tablets 4 times a day) per day respectively.. Duration 4 weeks. Concurrent medication/care: Propoxyphene napsylate was permitted as rescue medication during the washout period (100mg up

	to four times daily). This was not permitted during the trial.. Indirectness: No indirectness Comments: Ibuprofen 1200mg and 2400mg are reported in separate groups but have been combined for this analysis for class effect as agreed in the protocol (n=61) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 4000mg/day given as 8 tablets (2 tablets 4 times a day) per day for 4 weeks. Duration 4 weeks. Concurrent medication/care: Propoxyphene napsylate was permitted as rescue medication during the washout period (100mg up to four times daily). This was not permitted during the trial.. Indirectness: No indirectness
Funding	Academic or government funding (Supported by a grant (AR-39250) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Hospital assessment questionnaire pain score at 4 weeks; Group 1: mean 0.325 (SD 0.86); n=122, Group 2: mean 0.33 (SD 0.75); n=60; Hospital assessment questionnaire pain score 0-3 Top=High is poor outcome; Comments: Reports mean change (95% confidence intervals). Calculated SD from this. Reported ibuprofen 1200mg = 0.30 (0.09 to 0.51). Reported ibuprofen 2400mg = 0.35 (0.13 to 0.57). Reported paracetamol: 0.33 (0.14 to 0.52). Baseline ibuprofen 1200mg = 1.50 (0.80). Baseline ibuprofen 2400mg = 1.61 (0.86). Baseline paracetamol = 1.46 (0.82).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: Hospital assessment questionnaire disability score at 4 weeks; Group 1: mean 0.1 (SD 0.41); n=122, Group 2: mean 0.08 (SD 0.32); n=60; Hospital assessment questionnaire disability score 0-3 Top=High is poor outcome; Comments: Reports mean change (95% confidence intervals). Calculated SD from this. Reported ibuprofen 1200mg = 0.08 (-0.01 to 0.16). Reported ibuprofen 2400mg = 0.11 (-0.02 to 0.23). Reported paracetamol: 0.08 (0.00 to 0.16). Baseline ibuprofen 1200mg = 0.91 (0.67). Baseline ibuprofen 2400mg = 0.85 (0.69). Baseline paracetamol = 0.86 (0.61).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal adverse events at 4 weeks; Group 1: 21/123, Group 2: 10/61; Comments: Ibuprofen 1200mg = 7, ibuprofen 2400mg = 14 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5 Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Oedema at 4 weeks; Group 1: 5/123, Group 2: 1/61; Comments: Ibuprofen 1200mg = 2, ibuprofen 2400mg = 3 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5 Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Renal adverse events at 4 weeks; Group 1: 0/123, Group 2: 1/61 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5 Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system adverse events at 4 weeks; Group 1: 11/123, Group 2: 4/61; Comments: Ibuprofen 1200mg = 7, ibuprofen 2400mg = 4 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, duration of disease, weight, sex, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ibuprofen 1200mg: noncompliance = 7, uncontrolled pain = 1, adverse events = 5. Ibuprofen 2400mg: noncompliance = 6, adverse events = 6.; Group 2 Number missing: 16, Reason: Noncompliance = 10, uncontrolled pain = 1, adverse events = 5	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Breivik 2010[37]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=199)
Countries and setting	Conducted in Denmark, Finland, Norway, Sweden; Setting: Outpatient follow up
Line of therapy	2nd line
Duration of study	Other: 24 week double blind phase, 30 day additional follow up period after this
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who fulfilled the American College of Rheumatology criteria for osteoarthritis with radiographic evidence of osteoarthritis in the hip and/or knee as defined by grade 2-4 of the Kellgren Lawrence scale
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People who had a clinical diagnosis of osteoarthritis of the hip and/or knee and fulfilled the American College of Rheumatology criteria for osteoarthritis, with pain from the relevant joint for at least 1 year prior to enrollment. They had radiographic evidence of osteoarthritis of the hip and/or knee, and were taking NSAIDs or coxibs for their osteoarthritis pain for at least 1 month prior to the Screening Visit, at a stable frequency and dose and at least half the maximum allowed daily dose which gives an anti-inflammatory effect. They continued to experience at least moderate pain when walking on a flat surface, in spite of treatment with NSAIDs or coxibs and were willing to continue their treatment with NSAID or coxibs at a stable frequency and dose, until the end of the double blind phase. Those using intermittently low-potent opioids (e.g. tramadol, low dose codeine) were willing to discontinue this regimen for the screening visit until the completion or discontinuation visit and take paracetamol tablets provided by the Sponsor as intermittent analgesic rescue. Finally those who were receiving transcutaneous nerve stimulation or biofeedback prior to study entry were willing to discontinue this therapy for the duration of the study.
Exclusion criteria	People who had been treated with strong opioid analgesics (e.g. morphine, oxycodone, methadone, fentanyl-patch); people treated regularly with weak opioid analgesics such as tramadol, or codeine, for longer than three weeks prior to the screening visit; if any intermittent, short-term treatment with weak opioids could not be discontinued for the duration of the study; if they had a history of other chronic conditions for which they required frequent analgesic therapy (e.g. headaches, migraine, gout); were scheduled for any major surgery that would fall within the screening phase or the double blind phase of the study; if transcutaneous nerve stimulation or biofeedback prior to enrollment could not be discontinued for the

	duration of the study; if the investigator deemed that the person had any contraindication to treatment with opioid medication, such as history of alcohol or substance abuse; if the person had any other clinically significant disease or any reduced organ function; if the person was using antidepressants, antiepileptic drugs, steroids, hypnotics (that may increase respiratory depression of buprenorphine); if the person, or any close relatives, had long-QT syndrome, were on anti-arrhythmic medication (class 1A or class III) or had any unstable or symptomatic cardiac abnormality
Recruitment/selection of patients	People were recruited mainly from pain clinics and rheumatology clinics as well as through newspaper advertisement
Age, gender and ethnicity	Age - Mean (SD): 62.9 (9.5). Gender (M:F): 63:136. Ethnicity: All people were Caucasian
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip and/or knee).
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: At least 1 year.
Indirectness of population	No indirectness
Interventions	(n=100) Intervention 1: Opioids (topical - systemic) - Opioids (topical systemic). Buprenorphine patch, initial dose 5 microgram/hour uptitrated to a maximum dose of 20 micrograms/hour (only after a minimum of three days treatment of any given dose of buprenorphine).. Duration 24 weeks. Concurrent medication/care: Continuation of their current NSAID or coxib analgesic regimen at a stable frequency and dose. Rescue analgesia was provided as paracetamol tablets 0.5g for breakthrough osteoarthritis pain until the end of the double blind phase, up to 4g allowed daily.. Indirectness: No indirectness (n=99) Intervention 2: Placebo. Matching placebo patch. Duration 24 weeks. Concurrent medication/care: Continuation of their current NSAID or coxib analgesic regimen at a stable frequency and dose. Rescue analgesia was provided as paracetamol tablets 0.5g for breakthrough osteoarthritis pain until the end of the double blind phase, up to 4g allowed daily.. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Norpharma A/S)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OPIOIDS (TOPICAL SYSTEMIC) versus PLACEBO	

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC OA index of pain at 24 weeks; Group 1: mean -3.2 (SD 3.8); n=95, Group 2: mean -2.3 (SD 3.7); n=99; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline buprenorphine: 10.8 (2.6). Baseline placebo: 10.6 (2.8).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, weight and baseline values of outcomes; Group 1 Number missing: 54, Reason: 31 due to adverse events, 7 due to lack of efficacy, 4 without a reason, 5 due to protocol violations. However, the paper states a total of 54 people withdrew.; Group 2 Number missing: 34, Reason: 2 due to adverse events, 12 due to lack of efficacy, 5 without a reason, 6 due to protocol violations. However, the paper states a total of 34 people withdrew.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC OA index for functional ability at 24 weeks; Group 1: mean -10 (SD 11.7); n=94, Group 2: mean -6.5 (SD 11.4); n=96; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline buprenorphine: 37.4 (9.0). Baseline placebo: 35.1 (9.8).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, weight and baseline values of outcomes; Group 1 Number missing: 54, Reason: 31 due to adverse events, 7 due to lack of efficacy, 4 without a reason, 5 due to protocol violations. However, the paper states a total of 54 people withdrew.; Group 2 Number missing: 34, Reason: 2 due to adverse events, 12 due to lack of efficacy, 5 without a reason, 6 due to protocol violations. However, the paper states a total of 34 people withdrew.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastrointestinal disorders (including abdominal discomfort, constipation, nausea and vomiting) at 24 weeks; Group 1: 57/100, Group 2: 25/99; Comments: Buprenorphine: Gastro-intestinal disorders (total) = 57, abdominal discomfort = 1, constipation = 24, nausea = 37, vomiting = 16. Placebo: Gastro-intestinal disorders (total) = 25, constipation = 5, nausea = 10, vomiting = 2
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, weight and baseline values of outcomes; Group 1 Number missing: 54, Reason: 31 due to adverse events, 7 due to lack of efficacy, 4 without a reason, 5 due to protocol violations. However, the paper states a total of 54 people withdrew.; Group 2 Number missing: 34, Reason: 2 due to adverse events, 12 due to lack of efficacy, 5 without a reason, 6 due to protocol violations. However, the paper states a total of 34 people withdrew.

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Nervous system disorders (including dizziness, headache, somnolence) at 24 weeks; Group 1: 45/100, Group 2: 18/99; Comments: Buprenorphine: Nervous system disorders (total) = 45, dizziness = 25, headache = 7, somnolence = 4. Placebo: Nervous system disorders (total) = 18, dizziness = 9, headache = 6, somnolence = 0
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, weight and baseline values of outcomes; Group 1 Number missing: 54, Reason: 31 due to adverse events, 7 due to lack of efficacy, 4 without a reason, 5 due to protocol violations. However, the paper states a total of 54 people withdrew.; Group 2 Number missing: 34, Reason: 2 due to adverse events, 12 due to lack of efficacy,

5 without a reason, 6 due to protocol violations. However, the paper states a total of 34 people withdrew.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Burch 2007[39]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=646)
Countries and setting	Conducted in Canada, France, Romania, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Pain associated with osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People from 40-80 years old with pain due to osteoarthritis of the knee who were taking NSAIDs, COX-2 inhibitors, or tramadol on a regular basis for osteoarthritis pain during the 30 days that preceded enrollment
Exclusion criteria	A diagnosis of arthritis other than osteoarthritis; a history of injury or procedure that would interfere with assessment of pain in the knee; current or prior substance abuse or dependency; people who had been treated with any drug that reduces their seizure threshold within 3 weeks prior to screening
Recruitment/selection of patients	Multicenter study consisting of an open label phase and a double blind phase in 108 outpatient clinics
Age, gender and ethnicity	Age - Mean (SD): 62 (9). Gender (M:F): 238:408. Ethnicity: Asian = 2, Black = 33, Caucasian = 565, Hispanic = 43, Other = 3
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: not stated.
Indirectness of population	No indirectness
Interventions	(n=432) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol contramid once daily starting at 100mg and being increased in 100mg increments up to 300mg prior to the maintenance period. Duration 12 weeks. Concurrent medication/care: Paracetamol could be used during the titration phase. During the study, people were not permitted to take pain medication other than the study drug, with the exception of short-acting analgesics for acute pain other than that due to osteoarthritis. They could only be taken for a maximum of 3 consecutive days and not within 3 days of an

	assessment visit. Indirectness: No indirectness (n=214) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: Paracetamol could be used during the titration phase. During the study, people were not permitted to take pain medication other than the study drug, with the exception of short-acting analgesics for acute pain other than that due to osteoarthritis. They could only be taken for a maximum of 3 consecutive days and not within 3 days of an assessment visit. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Pain intensity numeric rating scale at 12 weeks; Group 1: mean 3.03 (SD 2.12); n=393, Group 2: mean 2.29 (SD 1.97); n=196; Numeric rating scale 0-10 Top=High is poor outcome; Comments: Baseline tramadol: 7.2 (1.6). Baseline placebo: 7.2 (1.6). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, ethnicity, BMI and baseline numeric rating scale pain intensity; Group 1 Number missing: 106, Reason: Adverse events = 44, treatment failure = 34, patient request = 23, other = 5; Group 2 Number missing: 49, Reason: Adverse event = 11, treatment failure = 22, patient request = 6, other = 10	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Cahlin 2011[40]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=59)
Countries and setting	Conducted in Sweden; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People referred with a diagnosis of temporomandibular disorder with the presence of osteoarthritis in one or both temporomandibular joints, fulfilling the Research Diagnostic criteria axis 1, IIIb i.e. pain and tenderness in the joint capsule and/or synovial lining of the temporomandibular joints and course crepitus or degenerative condition on tomograms
Stratum	TMJ
Subgroup analysis within study	Not applicable
Inclusion criteria	People referred with a diagnosis of temporomandibular disorder with the presence of osteoarthritis in one or both temporomandibular joints, fulfilling the Research Diagnostic criteria axis 1, IIIb i.e. pain and tenderness in the joint capsule and/or synovial lining of the temporomandibular joints and course crepitus or degenerative condition on tomograms. People had to be at least 18 years old and able to understand and follow instructions in Swedish.
Exclusion criteria	Unwillingness to participate; pharmacological treatment for any other painful condition; allergy/hypersensitivity to glucosamine or shellfish; pregnant/nursing
Recruitment/selection of patients	New consecutive patients.
Age, gender and ethnicity	Age - Mean (SD): 59 (12). Gender (M:F): 8:51. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Mixed 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): TMJ
Extra comments	Severity: Not stated Duration of symptoms: 2 to >24 months - the majority had symptoms for more than 24 months (median 13-24 months). The glucosamine in this study is not stated to have been quality assessed. It was made in the hospital with no involvement with pharmaceutical companies
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine capsules 400mg three times a day.

	Duration 6 weeks. Concurrent medication/care: Paracetamol 1000mg could be used for rescue medication. Indirectness: No indirectness (n=29) Intervention 2: Placebo. Placebo three times a day. Duration 6 weeks. Concurrent medication/care: Paracetamol 1000mg could be used for rescue medication. Indirectness: No indirectness
Funding	Academic or government funding (Supported by grants from the Research and Development Council of Göteborg and Södra Bohuslän)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for TMJ: Pain (visual analogue scale) at 6 weeks; Group 1: mean 38.7 (SD 28.8); n=30, Group 2: mean 36.8 (SD 20.8); n=29; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 48.1 (25.8). Baseline placebo: 42.5 (27.8). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, symptom duration, degenerative radiographic changes, radiographic findings, and previous treatment; Group 1 Number missing: 8, Reason: 8 discontinued. 5 due to gastrointestinal adverse events. 3 due to other adverse events or no reason given.; Group 2 Number missing: 2, Reason: 2 discontinued due to other adverse events or no reason given Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for TMJ: Gastrointestinal adverse events at 6 weeks; Group 1: 10/30, Group 2: 3/29 Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, symptom duration, degenerative radiographic changes, radiographic findings, and previous treatment; Group 1 Number missing: 8, Reason: 8 discontinued. 5 due to gastrointestinal adverse events. 3 due to other adverse events or no reason given.; Group 2 Number missing: 2, Reason: 2 discontinued due to other adverse events or no reason given	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Caldwell 2002[41]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=295)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis and grade II-IV radiographic evidence of osteoarthritis of the hip and/or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People had to be of at least 40 years of age and have both a clinical diagnosis and grade II-IV radiographic evidence of osteoarthritis of the hip and/or knee; have had prior suboptimal analgesic response to treatment with NSAIDs and paracetamol or had previously received intermittent opioid analgesic therapy; and have a baseline visual analogue scale pain intensity score of ≥40mm in the index joint
Exclusion criteria	People with serious concomitant disease; chronic conditions that might interfere with the assessment of pain and other symptoms of ostearthritis; prior disease at the index joint; surgery or the likelihood of requiring a surgical procedure of the index joint(s) during the trial; diseases other than osteoarthritis not well managed with treatment; weight ≤100lbs; oral, intramuscular, intravenous, intra-articular, or soft tissue administration of steroids within 1 month of study drug administration (two months, if at index knee or hip joint); intra-articular viscosupplementation (in the index joint) within 6 months of trial treatment; opioid therapy for longer than 3 weeks prior to baseline; any history of substance abuse within 2 years prior to screening; history of clinically significant intolerance to opioids or any known hypersensitivity to morphine or other opioid analgesics
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 62.4 (10.5). Gender (M:F): 111:184. Ethnicity: White = 249, Black = 40, Other = 6
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).

Extra comments	Severity: Radiographic grade II-IV Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=222) Intervention 1: Strong opioids (oral) - Morphine. Three groups that have been merged for this analysis: Avinza (long acting morphine sulphate) QAM (30mg in the morning, with a placebo version of morphine sulphate contin, and two placebo tablets in the evening); Avinza QPM (two placebo tablets in the morning, 30mg in the evening with a placebo) and morphine sulphate contin (15mg in the morning with a placebo version of Avinza, 15mg in the evening with a placebo version of Avinza). The Avinza dose could be increased if the optimal pain relief was not achieved. . Duration 4 weeks. Concurrent medication/care: The use of analgesic preparations other than the cardiovascular prophlyactic doses of aspirin (up to 325mg/day) and paracetamol for non-osteoarthritis symptomatology (up to 2000mg/day for a maximum of 3 consecutive days) was prohibited. Paracetamol had to be stopped 24 hours prior to efficacy assessments. Inhaled and topical steroids were permitted for the treatment of respiratory and dermatological disorders respectively. . Indirectness: No indirectness (n=73) Intervention 2: Placebo. Matching placebo - two tablets twice a day. Duration 4 weeks. Concurrent medication/care: The use of analgesic preparations other than the cardiovascular prophlyactic doses of aspirin (up to 325mg/day) and paracetamol for non-osteoarthritis symptomatology (up to 2000mg/day for a maximum of 3 consecutive days) was prohibited. Paracetamol had to be stopped 24 hours prior to efficacy assessments. Inhaled and topical steroids were permitted for the treatment of respiratory and dermatological disorders respectively. . Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: MORPHINE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC osteoarthritis index pain at 4 weeks; Group 1: mean -20.6 (SD 29.4); n=136, Group 2: mean -6.48 (SD 31.1); n=50; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports the scale was measured in 0-500, but reports percentage change from baseline (effectively WOMAC pain 0-100 - least mean square and standard error). Standard error converted to standard deviation and morphine groups combined. Reported Avinza AM = -19.3 (4.3). Reported Avinza PM = -22.2 (4.3). Reported MSC = -20.5 (4.3). Reported placebo = -6.48 (4.4). Baseline Avinza AM (0-500): 317 (102). Baseline Avinza PM = 326 (99.7). Baseline MSC: 322 (109). Baseline placebo: 317 (102).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, and baseline values

of outcomes; Group 1 Number missing: 88, Reason: Avinza QAM: 17 adverse events, 9 lack of efficacy, 1 other. Avinza QPM: 18 adverse events, 12 lack of efficacy, 3 other. MSC: 18 adverse events, 8 lack of efficacy, 2 other.; Group 2 Number missing: 23, Reason: 5 adverse events, 14 lack of efficacy, 4 other Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Other: WOMAC physical function at 4 weeks; Group 1: mean -198.3 (SD 274.6); n=134, Group 2: mean -96.7 (SD 304.1); n=50; WOMAC physical function 0-1700 Top=High is poor outcome; Comments: Reports the scale was measured in 0-1700 (least mean square and standard error). Standard error converted to standard deviation and morphine groups combined. Reported Avinza AM = -207 (40.7). Reported Avinza PM = -204 (42.6). Reported MSC = -181 (40.1). Reported placebo = -96.7 (43.0). Baseline Avinza AM: 1102 (328). Baseline Avinza PM = 1126 (328). Baseline MSC: 1112 (343). Baseline placebo: 1121 (313). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, and baseline values of outcomes; Group 1 Number missing: 88, Reason: Avinza QAM: 17 adverse events, 9 lack of efficacy, 1 other. Avinza QPM: 18 adverse events, 12 lack of efficacy, 3 other. MSC: 18 adverse events, 8 lack of efficacy, 2 other.; Group 2 Number missing: 23, Reason: 5 adverse events, 14 lack of efficacy, 4 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Caroit 1976[42]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=9)
Countries and setting	Conducted in France; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip based on classical criteria, associated with the presence of positive radiological signs
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People with painful osteoarthritis of the hip who responded favorably in the preceding year to anti-inflammatory treatment with non-steroidal drugs. Includes people with primary or secondary osteoarthritis.
Exclusion criteria	Rapidly destructive cases of osteoarthritis of the hip; cases requiring surgical treatment; pregnant women; subjects having a classical contra-indication to anti-inflammatory medication such as a history of gastroduodenal ulcers, renal or hepatic insufficiency, or severe haematological disease
Recruitment/selection of patients	Includes people who responded favourably to anti-inflammatory treatment in the past year
Age, gender and ethnicity	Age - Mean (range): 63 (54-72). Gender (M:F): 5:4. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Not stated Duration of symptoms (mean): 9 years
Indirectness of population	No indirectness
Interventions	(n=9) Intervention 1: NSAIDs - Ketoprofen. Ketoprofen 50mg three times a day. Duration 2 weeks. Concurrent medication/care: All other anti-inflammatory drugs, whether given systemically or locally, were excluded throughout the duration of the trial. Analgesics were withdrawn except for aspirin in so far as it had been used before and provided that the treatment was continued at the same dosage during the two treatment periods, A and B.. Indirectness: No indirectness (n=9) Intervention 2: Placebo. Placebo three times a day. Duration 2 weeks.

	Concurrent medication/care: All other anti-inflammatory drugs, whether given systemically or locally, were excluded throughout the duration of the trial. Analgesics were withdrawn except for aspirin in so far as it had been used before and provided that the treatment was continued at the same dosage during the two treatment periods, A and B.. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN versus PLACEBO Protocol outcome 1: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Hip: Abnormal findings for blood urea, alkaline phosphatases, transaminases, prothrombin determination or the blood count at 2 weeks; Group 1: 0/9, Group 2: 0/9 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex and duration of illness; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Case 2003[43] (Via 2003[198])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=82)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Unilaterally symptomatic idiopathic osteoarthritis of the knee based on radiographic and clinical enrollment criteria. The radiographic criteria consisted of the presence of radiographic osteoarthritis (modified Kellgren-Lawrence grade ≥1) in addition to evidence of medial compartment involvement, as evidence by possible or definite medial joint space narrowing or osteophytes. The clinical criteria included the presence of preenrollment ambulatory pain (defined as a visual analogue scale score of ≥30mm on the 100-mm scale corresponding to question 1 of the WOMAC pain section; moderate pain by a 5-point Likert scale; or increased pain (defined as an increase of ≥10mm by the visual analogue scale or ≥1 by the Likert scale) during a 2 week washout period following discontinuation of preexisting analgesic and/or anti-inflammatory osteoarthritis medications.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with unilateral symptomatic idiopathic osteoarthritis of the knee based on clinical and radiological criteria. In addition people had to be capable of independent ambulation without the aid of a cane or walker and had to fall within 1 SD of weight for their height and age according to the Metropolitan Life Insurance Company tables.
Exclusion criteria	Prior intolerance to either of the study medications or a history of an NSAID allergy or intolerance; functional class I or IV; history of peptic ulcer disease or of significant other gastrointestinal disease; significant hepatic abnormality (aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase level >20% above the upper limit of normal); renal insufficiency (creatinine level >1.2mg/dL [>106 micromol/L]); haematologic disease (haemoglobin level >1.2 mg/dL, white blood cell count <4.0x10^3/microL, or platelet count <120x10^3/microL); presence of joint disease other than osteoarthritis; presence of joint replacements in the lower extremity; use of substances that might interfere with pain perception (tranquilizers, hypnotic agents, or excessive alcohol intake), and anticoagulation therapy; the use of nonstudy pain medications during the trial

Recruitment/selection of patients	An option for the inclusion criteria was that people had an increase in pain during the washout period.
Age, gender and ethnicity	Age - Mean (SD): 62.2 (9.6). Gender (M:F): 41:41. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence mean grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=29) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg (two 500mg tablets) four times daily with matching placebo twice daily. Duration 12 weeks. Concurrent medication/care: Use of nonstudy pain medications during the trial was prohibited. Indirectness: No indirectness (n=25) Intervention 2: NSAIDs - Diclofenac. Diclofenac sodium 75mg twice daily with matching placebo two tablets four times a day. Duration 12 weeks. Concurrent medication/care: Use of nonstudy pain medications during the trial was prohibited. Indirectness: No indirectness (n=28) Intervention 3: Placebo. Placebo (paracetamol placebo) two tablets four times daily and placebo (diclofenac placebo) 1 tablet twice daily. Duration 12 weeks. Concurrent medication/care: Use of nonstudy pain medications during the trial was prohibited. Indirectness: No indirectness
Funding	Academic or government funding (This study was supported in part by a Specialised Center of Research osteoarthritis grant AR39239 from the National Institutes of Health, Bethesda, Md; and an intramural development grant from the Rush Arthritis and Orthopaedics Institute, Chicago, Ill (Dr Case). Mr Baliunas was the recipient of a Dean's Summer Research Fellowship from Rush Medical College, Chicago.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -23.8 (SD 83.2); n=29, Group 2: mean -15.3 (SD 98.7); n=28; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline paracetamol: 210.8 (86.3). Baseline placebo: 198.6 (110.9).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study

knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 7, Reason: 2 withdrew between weeks 0-2. 5 withdrew between weeks 2-12. This was due to inefficacy in 5 of the 7.; Group 2 Number missing: 9, Reason: 2 withdrew between weeks 0-2. 7 withdrew between weeks 2-12. 4 were due to inefficacy. 2 due to non-knee pain. 3 due to other reasons.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 12 weeks; Group 1: mean -41.8 (SD 205.6); n=29, Group 2: mean -85.6 (SD 223.2); n=28; WOMAC function subscale 0-1700 Top=High is poor outcome; Comments: Baseline paracetamol: 657.0 (262.5). Baseline placebo: 697.1 (375.2).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 7, Reason: 2 withdrew between weeks 0-2. 5 withdrew between weeks 2-12. This was due to inefficacy in 5 of the 7.; Group 2 Number missing: 9, Reason: 2 withdrew between weeks 0-2. 7 withdrew between weeks 2-12. 4 were due to inefficacy. 2 due to non-knee pain. 3 due to other reasons.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -53.9 (SD 79.3); n=25, Group 2: mean -23.8 (SD 83.2); n=29; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline diclofenac: 199.8 (101.5). Baseline paracetamol: 210.8 (86.3).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 5, Reason: 3 withdrew between weeks 0-2. 2 withdrew between weeks 2 and 12. This was due to adverse events in 3 of the 5.; Group 2 Number missing: 7, Reason: 2 withdrew between weeks 0-2. 5 withdrew between weeks 2-12. This was due to inefficacy in 5 of the 7.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 12 weeks; Group 1: mean -163 (SD 201.5); n=25, Group 2: mean -41.8 (SD 205.6); n=29; WOMAC function subscale 0-1700 Top=High is poor outcome; Comments: Baseline diclofenac: 669.3 (371.6). Baseline paracetamol: 657.0 (262.5).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 5, Reason: 3 withdrew between weeks 0-2. 2 withdrew between weeks 2 and 12. This was due to adverse events in 3 of the 5.; Group 2 Number missing: 7, Reason: 2 withdrew between weeks 0-2. 5 withdrew between weeks 2-12. This was due to inefficacy in 5 of the 7.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -53.9 (SD 79.3); n=25, Group 2: mean -15.3 (SD 98.7); n=28; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline diclofenac: 199.8 (101.5). Baseline placebo: 198.6 (110.9).

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 5, Reason: 3 withdrew between weeks 0-2. 2 withdrew between weeks 2 and 12. This was due to adverse events in 3 of the 5.; Group 2 Number missing: 9, Reason: 2 withdrew between weeks 0-2. 7 withdrew between weeks 2-12. 4 were due to inefficacy. 2 due to non-knee pain. 3 due to other reasons. Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC function subscale at 12 weeks; Group 1: mean -163 (SD 201.5); n=25, Group 2: mean -85.6 (SD 223.2); n=28; WOMAC function subscale 0-1700 Top=High is poor outcome; Comments: Baseline diclofenac: 669.3 (371.6). Baseline placebo: 697.1 (375.2). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, study knee, prestudy pain medications, prestudy WOMAC variables and radiographic features; Group 1 Number missing: 5, Reason: 3 withdrew between weeks 0-2. 2 withdrew between weeks 2 and 12. This was due to adverse events in 3 of the 5.; Group 2 Number missing: 9, Reason: 2 withdrew between weeks 0-2. 7 withdrew between weeks 2-12. 4 were due to inefficacy. 2 due to non-knee pain. 3 due to other reasons.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Chan 2010[44] (Kellner 2012[101])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=4484)
Countries and setting	Conducted in Belgium, Brazil, Canada, China, Colombia, Costa Rica, Croatia, Czech Republic, Ecuador, Estonia, France, Germany, Greece, Guatemala, Hong Kong (China), India, Latvia, Lithuania, Netherlands, Panama, Peru, Portugal, Russia, Serbia, Singapore, South Africa, South Korea, Spain, Sweden, Taiwan, Ukraine, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks (6 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a clinical diagnosis of osteoarthritis or rheumatoid arthritis who were expected to need regular NSAID treatment for at least 6 months
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a clinical diagnosis of osteoarthritis or rheumatoid arthritis were eligible if they were expected to need regular NSAID treatment for at least 6 months. Enrolment included those aged 60 years or older with or without a history of gastroduodenal ulceration or gastrointestinal haemorrhage. People aged between 18 years and 59 years were enrolled if they had a documented history of gastroduodenal ulceration or gastroduodenal haemorrhage more than 90 days before screening. People also had to test negative for Helicobacter pylori at the screening visit or have confirmed eradication of the infection at a rescreening visit.
Exclusion criteria	Concomitant use of antiplatelet or anticoagulant drugs; ischaemic heart disease; heart failure; peripheral arterial disease; cerebrovascular disease; gastrointestinal haemorrhage or active gastroduodenal ulceration less than 90 days before screening; inflammatory bowel disease; gastric surgery besides a patch repair; erosive oesophagitis; gastric-outlet obstruction; active malignant disease; alcohol and substance misuse; allergy to diclofenac, celecoxib, omeprazole or sulphonamides, serum alanine transaminase or aspartate transaminase concentrations more than 1.5 times and serum creatinine concentration more than 1.2 times the upper limit of normal (according to the central laboratory definition), and a haemoglobin concentration lower than 115g/L.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 65 (7.7). Gender (M:F): 814:3670. Ethnicity: White = 2450, Black = 106, Asian = 610, Hispanic = 926, Other = 392
Further population details	1. Age: Mixed (Based on range: 25-93). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Type not stated).
Extra comments	Severity: Not stated Duration of symptoms: Not stated. Results were supplemented with those from clinicaltrials.gov. NCT number:

	NCT00141102. (https://clinicaltrials.gov/ct2/show/results/NCT00141102?term=NCT00141102&draw=2&rank=1)
Indirectness of population	Serious indirectness: Includes people with rheumatoid arthritis, but less than 20% (17%)
Interventions	(n=2238) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg twice a day and diclofenac placebo plus omeprazole placebo. Duration 26 weeks. Concurrent medication/care: After randomisation people could take antacids or non-NSAID analgesic drugs, including paracetamol up to 4g per day and histamine-2-receptor antagonists no more than 3 days per week. Corticosteroids (prednisolone no more than 10mg daily), disease-modifying antirheumatic drugs, or biological treatments were only allowed if people had been taking a stable dose 12 or more weeks at randomisation. After randomisation, doses could be adjusted if clinically indicated for disease management; however, people were not allowed to start treatment with any of the agents during the study. NSAIDs other than study drugs, other antiulcer drugs, cytotoxic agents, lithium and iron supplements were prohibited.. Indirectness: No indirectness (n=2246) Intervention 2: NSAID and gastroprotection - NSAID and proton pump inhibitor. Diclofenac slow release 75mg twice a day plus omeprazole 20mg once a day and celecoxib placebo. Duration 26 weeks. Concurrent medication/care: After randomisation people could take antacids or non-NSAID analgesic drugs, including paracetamol up to 4g per day and histamine-2-receptor antagonists no more than 3 days per week. Corticosteroids (prednisolone no more than 10mg daily), disease-modifying antirheumatic drugs, or biological treatments were only allowed if people had been taking a stable dose 12 or more weeks at randomisation. After randomisation, doses could be adjusted if clinically indicated for disease management; however, people were not allowed to start treatment with any of the agents during the study. NSAIDs other than study drugs, other antiulcer drugs, cytotoxic agents, lithium and iron supplements were prohibited.. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus NSAID AND PROTON PUMP INHIBITOR

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Clinically significant events through the GI tract: including gastroduodenal (GD) hemorrhage; gastric outlet obstruction; GD, small or large bowel perforation; small or large bowel hemorrhage; clinically significant anemia of defined GI origin; acute GI hemorrhage of unknown origin, including presumed small bowel hemorrhage; clinically significant anemia of presumed occult GI origin including possible small bowel blood loss at 26 weeks; Group 1: 20/2238, Group 2: 81/2246; Comments: Celecoxib: total = 20. Gastroduodenal haemorrhage = 3, gastric outlet obstruction = 0, gastroduodenal, small bowel or large bowel perforation = 0, small bowel haemorrhage = 0, large bowel haemorrhage = 1, gastroduodenal ulcer or erosions = 5, early gastric cancer = 0, lower GI bleeding = 0, lower GI ulcer or erosions = 0, acute GI haemorrhage of unknown origin including presumed small-bowel haemorrhage = 1, clinically significant anaemia of presumed occult GI origin including possible small-bowel blood loss = 10. Diclofenac: total = 81. Gastroduodenal haemorrhage = 3, gastric outlet obstruction = 0, gastroduodenal, small bowel or large bowel perforation = 0, small bowel haemorrhage = 0, large bowel haemorrhage = 1,

gastroduodenal ulcer or erosions = 20, early gastric cancer = 1, lower GI bleeding = 1, lower GI ulcer or erosions = 2, acute GI haemorrhage of unknown origin including presumed small-bowel haemorrhage = 0, clinically significant anaemia of presumed occult GI origin including possible small-bowel blood loss = 53. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, osteoarthritis diagnosis, age, ethnicity, region of origin, haemoglobin, haematocrit, history of gastroduodenal ulcer or ulcer bleeding, previous Helicobacter pylori infection, comorbidity; Group 1 Number missing: 497, Reason: 15 did not receive treatment. 233 adverse events. 249 other.; Group 2 Number missing: 590, Reason: 9 did not receive treatment. 305 adverse events. 276 other.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Chappell 2009[46]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=231)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: OSteoarthritis of the knee as defined by the American College of Rheumatology clinical and radiographic criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients of at least 40 years of age with osteoarthritis of the knee with pain for ≥14 days of each month for 3 months before study entry, with a mean score ≥4 on the 24-hour average pain score (0-10) using the average of daily ratings from Visit 1 to Visit 2. People with bilateral osteoarthritis were required to identify an index knee on which to base ratings throughout the study. People had to agree to maintain their usual activity level throughout the course of the study.
Exclusion criteria	Body mass index >40kg/m²; a confounding painful condition that would interfere with assessment of the index joint; a diagnosis of inflammatory arthritis or an autoimmune disorder; or if they had received invasive therapies to the knee in the past 3 months, knee arthroscopy in the index knee within the past year, or joint replacement of the index knee at anytime; people who had a prior synovial fluid analysis indicative of a diagnosis other than osteoarthritis; people who were nonambulatory; people who required the use of crutches or a walker; people with psychiatric disorders including major depressive disorder (as identified using the Mini International Neuropsychiatric Interview); previous exposure to duloxetine; women who were pregnant or breastfeeding; a history of substance abuse or dependence; positive urine drug screen for any substance of abuse; existence of any serious medical or psychiatric condition that could compromise participation in the study; a history of recurrent seizures; uncontrolled narrow-angle glaucoma; acute liver injury or severe cirrhosis; known hypersensitivity to duloxetine or any of the inactive ingredients; frequent or severe allergic reactions to multiple medications.
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 62.3 (9.5). Gender (M:F): 80:151. Ethnicity: Caucasian = 194, African = 12, Hispanic = 19, Other = 6

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 7.0 (7.8) years .
Indirectness of population	No indirectness
Interventions	(n=111) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 30mg once daily increased to 60mg once daily after 2 weeks. Then continued for 5 weeks. At this point all people taking duloxetine were randomly reassigned to duloxetine 60mg once daily or duloxetine 120mg once daily. The dose was continued for 6 weeks and then underwent a 2 week tapering period (people on 120mg of duloxetine were reduced to 60mg for one week, then 30mg for 1 week, then discontinued. People on 60mg of duloxetine were reduced to 30mg for 1 week, then discontinued).. Duration 13 weeks. Concurrent medication/care: People who entered the trial taking an NSAID or paracetamol were allowed to continue taking the drug(s) during the study. People were stratified according to whether they were NSAID/paracetamol users. People were not allowed to have their dose of NSAIDs or paracetamol increased over what they were taking at visit 1 but were allowed to have their dose decreased or discontinued. Any change in or initiation of medications during the study require consultation with the investigator. Episodic use (no more than 3 consecutive days and not to exceed 20 total days during the study) of short-acting analgesics was allowed for acute injury or surgery or for rescue from an osteoarthritis knee pain flare.. Indirectness: No indirectness (n=120) Intervention 2: Placebo. Matching placebo. Duration 13 weeks. Concurrent medication/care: People who entered the trial taking an NSAID or paracetamol were allowed to continue taking the drug(s) during the study. People were stratified according to whether they were NSAID/paracetamol users. People were not allowed to have their dose of NSAIDs or paracetamol increased over what they were taking at visit 1 but were allowed to have their dose decreased or discontinued. Any change in or initiation of medications during the study require consultation with the investigator. Episodic use (no more than 3 consecutive days and not to exceed 20 total days during the study) of short-acting analgesics was allowed for acute injury or surgery or for rescue from an osteoarthritis knee pain flare.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Eli Lilly and Company)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO	

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D (UK index) at 13 weeks; Group 1: mean 0.21 (SD 0.2); n=103, Group 2: mean 0.11 (SD 0.21); n=114; EQ-5D -0.11-1 Top=High is good outcome; Comments: Reports mean change score (SE). Reported duloxetine: 0.21 (0.02). Reported placebo: 0.11 (0.02)
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, height, weight, duration of osteoarthritis since diagnosis and pain since onset, and baseline values of some outcomes. It does not report baseline outcome values for EQ-5D or WOMAC pain subscale.; Group 1 Number missing: 34, Reason: 15 adverse events, 8 subject decision, 2 lack of efficacy, 3 physician decision, 4 lost to follow up, 2 protocol violation; Group 2 Number missing: 24, Reason: 7 adverse events, 9 subject decision, 3 lack of efficacy, 2 physician decision, 1 protocol violation, 1 entry criteria not met, 1 sponsor decision

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -4.64 (SD 3.62); n=107, Group 2: mean -3.24 (SD 3.79); n=117; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change score (SE). Reported duloxetine: -4.64 (0.35). Reported placebo: -3.24 (0.35)
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, height, weight, duration of osteoarthritis since diagnosis and pain since onset, and baseline values of some outcomes. It does not report baseline outcome values for EQ-5D or WOMAC pain subscale.; Group 1 Number missing: 34, Reason: 15 adverse events, 8 subject decision, 2 lack of efficacy, 3 physician decision, 4 lost to follow up, 2 protocol violation; Group 2 Number missing: 24, Reason: 7 adverse events, 9 subject decision, 3 lack of efficacy, 2 physician decision, 1 protocol violation, 1 entry criteria not met, 1 sponsor decision

Protocol outcome 3: Psychological distress at ≤3- or >3- months
- Actual outcome for Knee: Beck Depression Inventory-II at 13 weeks; Group 1: mean -1.29 (SD 3.25); n=77, Group 2: mean -1.06 (SD 3.53); n=96; Becks Depression Inventory-II total score 0-63 Top=High is poor outcome; Comments: Reports mean change scores and standard error. Reported duloxetine: -1.29 (0.37). Reported placebo: -1.06 (0.36). Baseline duloxetine: 5.5 (6.8). Baseline placebo: 5.6 (5.9).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, height, weight, duration of osteoarthritis since diagnosis and pain since onset, and baseline values of some outcomes. It does not report baseline outcome values for EQ-5D or WOMAC pain subscale.; Group 1 Number missing: 34, Reason: 15 adverse events, 8 subject decision, 2 lack of efficacy, 3 physician decision, 4 lost to follow up, 2 protocol violation; Group 2 Number missing: 24, Reason: 7 adverse events, 9 subject decision, 3 lack of efficacy, 2 physician decision, 1 protocol violation, 1 entry criteria not met, 1 sponsor decision
- Actual outcome for Knee: Hospital Anxiety and Depression Scale (HADS) - Anxiety subscale at 13 weeks; Group 1: mean -1.35 (SD 2.55); n=77, Group 2: mean -0.88 (SD 2.37); n=96; Hospital Anxiety and Depression Scale - Anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports means and standard errors. Reported duloxetine: -1.35 (0.27), reported placebo: -0.88 (0.26).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, height, weight, duration of osteoarthritis since diagnosis and pain since onset, and baseline values of some outcomes. It does not report baseline outcome values for EQ-5D or WOMAC

pain subscale.; Group 1 Number missing: 34, Reason: 15 adverse events, 8 subject decision, 2 lack of efficacy, 3 physician decision, 4 lost to follow up, 2 protocol violation; Group 2 Number missing: 24, Reason: 7 adverse events, 9 subject decision, 3 lack of efficacy, 2 physician decision, 1 protocol violation, 1 entry criteria not met, 1 sponsor decision	
Protocol outcomes not reported by the study	Physical function at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Chappell 2011[45]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=256)
Countries and setting	Conducted in Canada, Greece, Russia, Sweden, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who met the American College of Rheumatology clinical and radiographic criteria for the diagnosis of osteoarthritis of the knee with pain for ≥14 days per month during three consecutive months preceding study entry
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients ≥40 years of age with osteoarthritis of the knee who had pain severity of ≥4 on the 24-hour average pain severity scale (0 to 10), using the mean of daily ratings from the week preceding randomisation. People had to consent to maintain their usual activity level throughout the course of the study.
Exclusion criteria	Body mass index >40kg/m²; diagnosis of inflammatory arthritis or an autoimmune disorder; history of invasive therapies to the index knee during the past 3 months or joint replacement of the index knee at anytime; prior synovial fluid analysis indicative of a diagnosis other than osteoarthritis; nonambulatory or crutch- or walker-dependent; presence of psychiatric disorders, including major depressive disorder (identified using the Mini International Neuropsychiatric Interview); previous exposure to duloxetine; for female patients, existing pregnancy or breastfeeding; history of substance abuse or dependence; presence of serious medical condition; history of recurrent seizures, uncontrolled narrow-angle glaucoma, acute liver injury or severe cirrhosis; known hypersensitivity to duloxetine or any of the inactive ingredients; and frequent or severe allergic reactions to multiple medications
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 62.6 (9.0). Gender (M:F): 60:196. Ethnicity: Caucasian = 250, African American = 3, Other = 3
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: 5.9 (6.1) years. NCT00433290. Results are supported by information published on clinicaltrials.gov (https://clinicaltrials.gov/ct2/show/results/NCT00433290?term=nct00433290&draw=2&rank=1)
Indirectness of population	No indirectness

Interventions	(n=128) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 60-120mg once daily for 13 weeks. This included starting people on 30mg duloxetine once daily for 1 week, then uptitrating to 60mg once daily. After week 7 the dosage was increased to 120mg once daily for people reporting <30% pain reduction from baseline using Brief Pain Inventory 24-hour average pain rating with no tolerability concerns. These people remainted on 120mg once daily for the remainder of the study. At week 13 a 2 week taper phase took place reducing the dose for people on 60mg to 30mg for 1 week before discontinuation and for people on 120mg to 60mg for 1 week then 30mg for 1 week before discontinuation.. Duration 13 weeks. Concurrent medication/care: People who entered the trial taking an NSAID or paracetamol were allowed to continue taking it provided that the dosage was not increased during the study. Randomisation of people was stratified according to whether they were NSAID users. After randomisation, episodic use (≤3 consecutive days and was not to exceed 20 total days during the study) of short acting analgesics was allowed for acute injury or surgery or for rescue from osteoarthritis flare up pain. Any change in or initiation of prescription or over-the-counter medications during the study required consultation with the study site personnel. Use of antidepressants and anticonvulsants or other agents used for the treatment of chronic pain, except as described previously, was not allowed.. Indirectness: No indirectness (n=128) Intervention 2: Placebo. Matching placebo for 13 weeks. Duration 13 weeks. Concurrent medication/care: People who entered the trial taking an NSAID or paracetamol were allowed to continue taking it provided that the dosage was not increased during the study. Randomisation of people was stratified according to whether they were NSAID users. After randomisation, episodic use (≤3 consecutive days and was not to exceed 20 total days during the study) of short acting analgesics was allowed for acute injury or surgery or for rescue from osteoarthritis flare up pain. Any change in or initiation of prescription or over-the-counter medications during the study required consultation with the study site personnel. Use of antidepressants and anticonvulsants or other agents used for the treatment of chronic pain, except as described previously, was not allowed.. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Eli Lilly and Company, Indianapolis, IN, U.S.A.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D at 13 weeks; Group 1: mean 0.09 (SD 0.16); n=121, Group 2: mean 0.08 (SD 0.18); n=124; EQ-5D -0.11-1 Top=High is good outcome; Comments: Baseline duloxetine = 0.68 (0.16). Baseline placebo = 0.66 (0.16).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -4.27 (SD 3.3); n=123, Group 2: mean -3.49 (SD 3.89); n=127; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline duloxetine = 10.24 (2.47). Baseline placebo = 10.35 (2.66).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean -13.78 (SD 10.78); n=118, Group 2: mean -10.75 (SD 10.98); n=126; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline duloxetine = 35.05 (9.63). Baseline placebo = 36.82 (8.15).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nausea, constipation, upper abdominal pain, diarrhoea, dyspepsia, haemorrhoids at 13 weeks; Group 1: 9/128, Group 2: 2/128; Comments: Duloxetine: Nausea = 5, constipation = 1, upper abdominal pain = 0, diarrhoea = 1, dyspepsia = 1, haemorrhoids = 1. Placebo: Nausea = 0, constipation = 1, upper abdominal pain = 1, diarrhoea = 0, dyspepsia = 0, haemorrhoids = 0
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Adverse events reported are only those that led to discontinuation from the study; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.

Protocol outcome 5: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Atrial fibrillation, palpitations, supraventricular tachycardia at 13 weeks; Group 1: 2/128, Group 2: 1/128; Comments: Duloxetine: Atrial fibrillation = 0, palpitations = 1, supraventricular tachycardia = 1. Placebo: Atrial fibrillation = 1, palpitations = 0, supraventricular tachycardia = 0
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Adverse events reported are only those that led to discontinuation from the study; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.

Protocol outcome 6: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Acute pyelonephritis at 13 weeks; Group 1: 0/128, Group 2: 1/128

Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Adverse events reported are only those that led to discontinuation from the study; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other. Protocol outcome 7: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Insomnia, abnormal dreams, anxiety, memory impairment, sleep disorder at 13 weeks; Group 1: 5/128, Group 2: 2/128; Comments: Duloxetine: Insomnia = 1, abnormal dreams = 1, anxiety = 1, memory impairment = 1, sleep disorder = 1. Placebo: Insomnia = 2, abnormal dreams = 0, anxiety = 0, memory impairment = 0, sleep disorder = 0 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Adverse events reported are only those that led to discontinuation from the study; Baseline details: Reports age, ethnicity, gender, duration of osteoarthritis since diagnosis and pain since onset, NSAID use, and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 discontinued. 24 adverse events, 1 lack of efficacy, 10 other.; Group 2 Number missing: 17, Reason: 17 discontinued. 7 adverse events, 5 lack of efficacy, 5 other.	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Chindalore 2005[47]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=362)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate to severe pain in 1 or more hip or knee joints caused by osteoarthritis for at least 3 months
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women 18-70 years of age; women who were postmenopausal, physically incapable of childbearing, or practicing an acceptable method of birth control; ambulatory; moderate to severe pain in 1 or more hip or knee joints caused by osteoarthritis for at least 3 months before screening; moderate to severe pain in the hip or knee while taking 1 or more oral analgesic medications in the previous month; a pain intensity score ≥5 on an 11 point numeric scale; a mean daily diary overall pain intensity ≥5 during the last 2 days of the 4-7 day washout period and a confirmatory pain intensity level ≥5 measured at the clinic before randomisation; able to understand and cooperate with study procedures; and, no pain medications other than study drug during the 3 week pretreatment period, except for aspirin up to 325mg/day for cardiovascular prophylaxis
Exclusion criteria	A daily opioid dose equivalent to >20mg oxycodone for 2 or more days within the previous 4 weeks; administration of an opioid within 72 hours; body weight >300 lb or <100 lb; major surgery within 3 months or planned surgery during the study period; oral or parenteral corticosteroid therapy within 1 month; intra-articular injection of hyaluronic acid within 9 months; epidural or intrathecal infusion of any analgesic medication within 1 month; females who were pregnant or breast feeding; history of severe hepatic or renal impairement; acute hepatitis; known allergy or significant reaction to any of the study medications; severe impairment of pulmonary function, hypercarbia, hypoxia, significant chronic obstructive airways disease or cor pulmonale, acute or severe bronchial asthma, sleep apnoea syndrome or respiratory depression; paralytic ileus, acute abdomen (serious abdominal pain requiring emergency surgery) or delayed gastric emptying; chronic biliary tract disease, chronic pancreatitis, or inflammatory bowel disorders; untreated myxoedema, untreated hypothyroidism, elevated intracranial pressure, severe anaemia, adrenocortical insufficiency,

	hypotension or hypovolemia; monoaminoxidase inhibitors, tricyclic antidepressant drugs, serotonin reuptake inhibitors, glucosamine/chondroitin, or St. John's Wort within 4 weeks before receiving study medication (a constant dose for longer than 4 weeks was acceptable); high doses of sedatives, hypnotics, or tranquilizers; phenothiazines or other agents that compromise vasomotor tone; history of alcohol or drug abuse; previous administration of Oxytrex; participation in another investigational drug trial or therapeutic trial within 30 days of the screening visit; and analgesic medication (other than paracetamol, up to five 500mg caplets per day) during the 4-7 day washout period before randomisation
Recruitment/selection of patients	Conducted over 37 centers
Age, gender and ethnicity	Age - Mean (range): 54.3 (21-70). Gender (M:F): 111:251. Ethnicity: 80.8% White, 13.3% Black, 4.4% Hispanic, 0.8% Asian
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Moderate to severe Duration of symptoms: Not explicitly stated. At least 3 months.
Indirectness of population	No indirectness
Interventions	(n=103) Intervention 1: Strong opioids (oral) - Oxycodone. Escalating dose of oxycodone starting at 2.5mg four times a day, increasing to 5mg four times a day after 3 days, increasing to 7.5mg four times a day after a week, increasing to 10mg four times a day after 2 weeks.. Duration 3 weeks. Concurrent medication/care: Aspirin up to 325mg/day was allowed for cardiovascular prophylaxis. Paracetamol up to five 500mg caplets per day was allowed during the washout period.. Indirectness: No indirectness (n=207) Intervention 2: Strong opioids (oral) - Oxycodone. Oxycodone and naltrexone either starting at 2.5mg oxycodone and 0.001mg naltrexone four times a day increasing up to 10mg oxycodone and the same dose of naltrexone four times a day at 3 weeks or starting at 5mg oxycodone and 0.001mg naltrexone twice a day increasing up to 20mg oxycodone and the same dose of naltrexone twice a day at 3 weeks. Duration 3 weeks. Concurrent medication/care: Aspirin up to 325mg/day was allowed for cardiovascular prophylaxis. Paracetamol up to five 500mg caplets per day was allowed during the washout period.. Indirectness: No indirectness Comments: Naltrexone is not an agreed intervention for this review and the combination does not fall into the remit of this review therefore is not included in the

	analysis. It is reported for completeness. (n=52) Intervention 3: Placebo. Matching placebo four times a day. Duration 3 weeks. Concurrent medication/care: Aspirin up to 325mg/day was allowed for cardiovascular prophylaxis. Paracetamol up to five 500mg caplets per day was allowed during the washout period.. Indirectness: No indirectness
Funding	Study funded by industry (Funded by Pain Therapeutics, Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYCODONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain intensity scores at 3 weeks; Group 1: mean 5.6 (SD 2.3); n=102, Group 2: mean 6.1 (SD 2.8); n=51; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline oxycodone: 7.4 (1.3). Baseline placebo: 7.7 (1.3).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline pain intensity scores. Otherwise reports overall values for gender and age.; Group 1 Number missing: 32, Reason: 2 inadequate pain relief, 29 adverse events, 1 other; Group 2 Number missing: 10, Reason: 7 inadequate pain relief, 0 adverse events, 3 other

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastrointestinal disorders (including upper abdominal pain, constipation, diarrhoea, dry mouth, nausea and vomiting) at 3 weeks; Group 1: 54/102, Group 2: 12/51; Comments: Oxycodone: Total = 54, upper abdominal pain = 4, constipation = 19, diarrhoea = 9, dry mouth = 9, nausea = 9, vomiting = 15. Placebo: Total = 12, upper abdominal pain = 2, constipation = 4, diarrhoea = 4, dry mouth = 0, nausea = 6, vomiting = 3
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline pain intensity scores. Otherwise reports overall values for gender and age.; Group 1 Number missing: 32, Reason: 2 inadequate pain relief, 29 adverse events, 1 other; Group 2 Number missing: 10, Reason: 7 inadequate pain relief, 0 adverse events, 3 other

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Nervous system disorders (including dizziness, headache and somnolence) at 3 weeks; Group 1: 49/102, Group 2: 14/51; Comments: Oxycodone: total = 49, dizziness = 26, headache = 24, somnolence = 21. Placebo: total = 14, dizziness = 0, headache = 12, somnolence = 3
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline pain intensity scores. Otherwise reports overall values for gender and age.; Group 1 Number missing: 32, Reason: 2 inadequate pain relief, 29 adverse events, 1 other; Group 2 Number missing: 10, Reason: 7 inadequate pain relief, 0 adverse events, 3 other

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months

Study	Chopra 2013[48]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=440)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a diagnosis of knee osteoarthritis as per modified American College of Rheumatology classification criteria (the lower age limit was 40 years) and pain (visual analogue scale at least 4cm in one or both knees while performing a weight-bearing activity) during the preceding 24 hours
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either sex in the age range 40-70 with a diagnosis of knee osteoarthritis as per modified American College of Rheumatology classification criteria (the lower age limit was 40 years) and pain (visual analogue scale at least 4cm in one or both knees while performing a weight-bearing activity) during the preceding 24 hours. Ambulant people who required frequent analgesics.
Exclusion criteria	Pregnant or lactating women or women with childbearing potential and not following adequate contraception; people with non-degenerative joint disorders; severe disabling arthritis (including wheelchair users) or a history of spine and lower limb surgery; people on medication likely to influence efficacy evaluation (except paracetamol rescue); people with a history of peptic ulcer bleed or recent active peptic ulcer; people with any unstable severe medical disease
Recruitment/selection of patients	People with chronic knee pain were screened in outpatient clinics and cost-free community arthritis camps
Age, gender and ethnicity	Age - Mean (SD): 55.7 (8.3). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 54.9 (51.9) months.
Indirectness of population	No indirectness

Interventions	(n=110) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Oral glucosamine sulphate (2g daily) given over three times during a day in equal divided doses. The glucosamine underwent quality assurance.. Duration 26 weeks. Concurrent medication/care: Paracetamol (500mg tablet) was provided for emergency analgesic use. Ongoing concomitant medication for concurrent chronic illness was permitted. People were not allowed treatment with any other alternative medicinal system (such as homeopathy, acupuncture or acupressure). People could continue their regular exercise and/or physiotherapy programme begun prior to the current trial, but were discouraged from starting any new activity during the trial. Physical therapy and local applications of pain relieving ointments/gels were not permitted. People were not prescribed any instructions or advice regarding diet or other life style change as per standard Ayuveda practice.. Indirectness: No indirectness (n=110) Intervention 2: NSAIDs - Celecoxib. Celecoxib 200mg daily administered over three times in the day in equally divided doses.. Duration 26 weeks. Concurrent medication/care: Paracetamol (500mg tablet) was provided for emergency analgesic use. Ongoing concomitant medication for concurrent chronic illness was permitted. People were not allowed treatment with any other alternative medicinal system (such as homeopathy, acupuncture or acupressure). People could continue their regular exercise and/or physiotherapy programme begun prior to the current trial, but were discouraged from starting any new activity during the trial. Physical therapy and local applications of pain relieving ointments/gels were not permitted. People were not prescribed any instructions or advice regarding diet or other life style change as per standard Ayuveda practice.. Indirectness: No indirectness (n=220) Intervention 3: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Ayuvedic preparations of medications: shunthi-guduchi formulations (SGC and SGCG) each containing amallaki (Emblica officinalis) in addition SGCG contained guggul (Boswelia serrata). Duration 26 weeks. Concurrent medication/care: Paracetamol (500mg tablet) was provided for emergency analgesic use. Ongoing concomitant medication for concurrent chronic illness was permitted. People were not allowed treatment with any other alternative medicinal system (such as homeopathy, acupuncture or acupressure). People could continue their regular exercise and/or physiotherapy programme begun prior to the current trial, but were discouraged from starting any new activity during the trial. Physical therapy and local applications of pain relieving ointments/gels were not permitted. People were not prescribed any instructions or advice regarding diet or other life style change as per

	standard Ayuveda practice.. Indirectness: No indirectness Comments: Ayuvedic preparations of medications were not interventions included in the protocol so is not included in the analysis. It is reported here for completeness.
Funding	Academic or government funding (This work was fully funded and supported by NMITLI Cell, Council of Scientific and Industrial Research, Government of India)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus CELECOXIB

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain at 24 weeks; Group 1: mean -2.72 (SD 3.32); n=110, Group 2: mean -6.93 (SD 3.13); n=110; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports change score and 95% confidence intervals. Reported glucosamine: -2.72 (-3.34, -2.10). Reported celecoxib: -1.90 (-2.48, -1.31). Baseline glucosamine: 9.33 (3.37). Baseline celecoxib: 9.43 (2.83).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, disease duration and baseline values of outcomes; Group 1 Number missing: 24, Reason: 5 adverse events, 1 lost to follow up, 10 unsatisfactory improvement, 4 protocol violations, 4 other reasons; Group 2 Number missing: 32, Reason: 4 adverse events, 8 lost to follow up, 9 unsatisfactory improvement, 6 protocol violation, 5 other reasons

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC difficulty (physical function subscale) at 24 weeks; Group 1: mean -8.12 (SD 11.1); n=110, Group 2: mean -6.93 (SD 10.2); n=110; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports change score and 95% confidence intervals. Reported glucosamine: -8.12 (-10.20, -6.04). Reported celecoxib: -6.93 (-8.85, -5.02). Baseline glucosamine: 32.03 (11.04). Baseline celecoxib: 34.30 (10.00).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, disease duration and baseline values of outcomes; Group 1 Number missing: 24, Reason: 5 adverse events, 1 lost to follow up, 10 unsatisfactory improvement, 4 protocol violations, 4 other reasons; Group 2 Number missing: 32, Reason: 4 adverse events, 8 lost to follow up, 9 unsatisfactory improvement, 6 protocol violation, 5 other reasons

Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Elevated SGPT (serum glutamic-pyruvic transaminase or alanine transaminase) at 24 weeks; Group 1: 4/108, Group 2: 2/105
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, disease duration and baseline values of outcomes; Group 1 Number missing: 24, Reason: 5 adverse events, 1 lost to follow up, 10 unsatisfactory improvement, 4 protocol violations, 4 other reasons; Group 2 Number missing: 32, Reason: 4 adverse events, 8 lost to follow up, 9 unsatisfactory improvement, 6 protocol violation, 5 other reasons

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Cibere 2004[49]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=137)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks (6 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee(s) according to the American College of Rheumatology diagnostic criteria and Kellgren-Lawrence grade of at least 2 on anteroposterior radiograph of the knee
Stratum	Knee:
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee(s) according to the American College of Rheumatology diagnostic criteria; Kellgren-Lawrence grade of at least 2 on anteroposterior radiograph of the knee; current daily use of glucosamine for at least 1 month; at least moderate improvement in knee pain since starting on glucosamine, measured on a 6-point scale of knee pain
Exclusion criteria	Chondroitin sulfate use within the previous 2 months; knee injection with hyaluronate in the previous 6 months or with corticosteroids in the previous 3 months; surgical procedure on either knee in the previous 3 months; narcotic analgesic use; uncontrolled medical condition or planned surgery that could interfere with follow up; baseline potassium >5.3 mEq/liter or baseline creatinine >120 mmol/liter
Recruitment/selection of patients	People were recruited through newspaper advertising and local posters. All people had to be using glucosamine before the study started (is a withdrawal trial).
Age, gender and ethnicity	Age - Range: 65 (40-88). Gender (M:F): 60:77. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms (median [range]): 3 (0-29) years
Indirectness of population	No indirectness
Interventions	(n=71) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 500mg tablets formulated as a potassium salt. The dosage was equivalent to 1500mg per day.. Duration 26 weeks. Concurrent medication/care: Rescue analgesic medications including paracetamol and NSAIDs

	were allowed and recorded by the patient in a daily diary. Other concomitant treatments, including chondroitin sulfate and intraarticular injections with corticosteroids or hyaluronic acid, were not allowed during the study.. Indirectness: No indirectness (n=66) Intervention 2: Placebo. Matching placebo three times a day. Duration 26 weeks. Concurrent medication/care: Rescue analgesic medications including paracetamol and NSAIDs were allowed and recorded by the patient in a daily diary. Other concomitant treatments, including chondroitin sulfate and intraarticular injections with corticosteroids or hyaluronic acid, were not allowed during the study.. Indirectness: No indirectness
Funding	Academic or government funding (Supported by grants from the Mary Pack Research Fund, Vancouver, British Columbia, Canada and by the Doris Alma Mary Anderson Fund for Geriatric Research, London, ON, Canada. Dr Cibere's work supported by a Canadian Institutes of Health Research Clinician Scientist Award and a Michael Smith Foundation for Health Research Postdoctoral Fellowship Award)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D utility score at 26 weeks; Group 1: mean -0.03 (SD 0.16); n=71, Group 2: mean -0.04 (SD 0.2); n=66; EQ-5D 0-1 Top=High is poor outcome; Comments: Baseline values not reported
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of glucosamine use, prestudy type of glucosamine, prestudy glucosamine dosage, duration of osteoarthritis, radiographic severity, baseline values for WOMAC, analgesic medication use. Does not report baseline values for EQ-5D.; Group 1 Number missing: 35, Reason: 1 lost to follow up. 2 discontinued due to concurrent illness. 32 developed a flare and so finished the study before 26 weeks.; Group 2 Number missing: 28, Reason: 28 developed a flare and so finished the study before 26 weeks.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain scale at 26 weeks; Group 1: mean -25 (SD 98); n=71, Group 2: mean -28 (SD 104); n=66; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline glucosamine (median [range]): 86 (2-289). Baseline placebo (median [range]): 86 (4-301).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of glucosamine use, prestudy type of glucosamine, prestudy glucosamine dosage, duration of osteoarthritis, radiographic severity, baseline values for WOMAC, analgesic medication use. Does not report baseline values for EQ-5D.; Group 1 Number missing: 35, Reason: 1 lost to follow up. 2 discontinued due to concurrent illness. 32 developed a flare and so finished the study before 26 weeks.; Group 2 Number missing: 28, Reason: 28 developed a flare and so finished the study

before 26 weeks. Protocol outcome 3: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC function scale at 26 weeks; Group 1: mean -58 (SD 270); n=71, Group 2: mean -63 (SD 318); n=66; WOMAC function subscale 0-1700 Top=High is poor outcome; Comments: Baseline glucosamine (median [range]): 268 (0-1376). Baseline placebo (median [range]): 294 (2-1240). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of glucosamine use, prestudy type of glucosamine, prestudy glucosamine dosage, duration of osteoarthritis, radiographic severity, baseline values for WOMAC, analgesic medication use. Does not report baseline values for EQ-5D.; Group 1 Number missing: 35, Reason: 1 lost to follow up. 2 discontinued due to concurrent illness. 32 developed a flare and so finished the study before 26 weeks.; Group 2 Number missing: 28, Reason: 28 developed a flare and so finished the study before 26 weeks. Protocol outcome 4: Osteoarthritis flare-ups at ≤3- or >3- months - Actual outcome for Knee: Osteoarthritis flare at 26 weeks; Group 1: 32/71, Group 2: 28/66; Comments: Flare defined as either the patient's perception of worsening of symptoms with a concomitant increase by at least 20mm in WOMAC pain on walking, or a significant worsening in the physician global assessment by at least 1 grade (1-5 scale) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of glucosamine use, prestudy type of glucosamine, prestudy glucosamine dosage, duration of osteoarthritis, radiographic severity, baseline values for WOMAC, analgesic medication use. Does not report baseline values for EQ-5D.; Group 1 Number missing: 3, Reason: 1 lost to follow up. 2 discontinued due to concurrent illness.; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Conaghan 2013[51]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1399)
Countries and setting	Conducted in Czech Republic, Germany, Poland, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary diagnosis of functional class I-III knee osteoarthritis meeting the American College of Rheumatology clinical classification criteria for knee osteoarthritis, defined by knee pain and at least four of the following: 1) age >50 years; 2) morning stiffness of <30 minute duration; 3) crepitus on active motion; 4) bony tenderness; 5) bony enlargement; 6) no palpable warmth of the synovium. People aged 18-45 years were eligible if they had radiological confirmation of osteoarthritis.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged >45 years with a primary diagnosis of functional class I-III knee osteoarthritis and meeting the American College of Rheumatology clinical classification criteria for knee osteoarthritis or people aged 18-45 years with radiological confirmation of osteoarthritis. People had to be able to identify a predominantly painful (index) knee with moderate pain in the index knee, defined by pain on walking on a flat surface rather as at least 4 for question 1 of the WOMAC (11-point numerical rating scale) and a total average WOMAC pain subscale score of <7 at the first and second baseline visits (separated by 2-5 days).
Exclusion criteria	Skin lesions or dermatological diseases in the treatment area; BMI >37kg/m²; symptomatic ipsilateral hip osteoarthritis or predominant patellofemoral knee osteoarthritis; inflammatory arthritis; non-index knee pain, or any other pain condition, requiring the chronic use of pain medication; contraindications for oral NSAID use (coagulopathy or concomitant use of anticoagulants, ischaemic heart disease, arterial/cerebrovascular disease, history of stroke or myocardial infarction, congestive heart failure, history of pancreatitis or peptic ulcers or inflammatory gastrointestinal disease); intraarticular injections of hyaluronic acid within 3 months before or during the study; use of oral, inhaled or parenteral corticosteroids within 2 months before or during the study; intraarticular injections of corticosteroids 1 month before or during the study.

Recruitment/selection of patients	Has two baseline visits with a washout period of at least 5 days or five times the half-life of any analgesic they were on, whichever was longest. Participants had to keep a WOMAC pain subscale score of <7 and a score of at least 4 for question 1 of the WOMAC to be included.
Age, gender and ethnicity	Age - Mean (range): 61.2 (24-90). Gender (M:F): 475:924. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Mixed 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=463) Intervention 1: NSAID gels (topical - local) - Other. Ketoprofen gel either 50 or 100mg in 2.2 or 4.4 grams respectively applied twice a day. Apply a set volume of gel (determined by an administration aid) to the index knee on both sides of the knee and the popliteal fossa, but avoid the patella (and any wounds). To be spread in gently (leaving no white residue) using the fingers. Then leave the skin to dry for at least 15 minutes before putting on clothes.. Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg up to four times a day was permitted for intermittent pain treatment, although not within 24 hours of the next study visit or between the baseline visits. People requiring ≥2 grams of paracetamol or other analgesic medication for longer than 3 consecutive days were considered treatment failures and withdrawn from the study.. Indirectness: No indirectness Comments: These two groups were combined due to class effect as agreed in the protocol (n=235) Intervention 2: NSAIDs - Celecoxib. Celecoxib 100mg orally twice a day. Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg up to four times a day was permitted for intermittent pain treatment, although not within 24 hours of the next study visit or between the baseline visits. People requiring ≥2 grams of paracetamol or other analgesic medication for longer than 3 consecutive days were considered treatment failures and withdrawn from the study.. Indirectness: No indirectness (n=473) Intervention 3: Placebo. 2.2g or 4.4g TDT 064/vehicle - the ketoprofen gel vehicle without the ketoprofen. Apply a set volume of gel (determined by an administration aid) to the index knee on both sides of the knee and the popliteal fossa, but avoid the patella (and any wounds). To be spread in gently (leaving no white residue) using the fingers. Then leave the skin to dry for at least 15 minutes before

	putting on clothes.. Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg up to four times a day was permitted for intermittent pain treatment, although not within 24 hours of the next study visit or between the baseline visits. People requiring ≥2 grams of paracetamol or other analgesic medication for longer than 3 consecutive days were considered treatment failures and withdrawn from the study.. Indirectness: No indirectness Comments: These two groups were combined due to class effect as agreed in the protocol (n=228) Intervention 4: Placebo. Placebo tablets twice a day. Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg up to four times a day was permitted for intermittent pain treatment, although not within 24 hours of the next study visit or between the baseline visits. People requiring ≥2 grams of paracetamol or other analgesic medication for longer than 3 consecutive days were considered treatment failures and withdrawn from the study.. Indirectness: No indirectness
Funding	Study funded by industry (Editorial assistance with the preparation of the manuscript was provided by Kirsteen Munn of Bollin Strategies Ltd, UK, and was funded by Pro Bono Bio Entrepeneur Ltd, UK)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus CELECOXIB

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -1.9 (SD 1.67); n=463, Group 2: mean -1.9 (SD 1.62); n=233; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Baseline ketoprofen 50mg: 4.7 (1.1). Baseline ketoprofen 100mg: 4.8 (0.9). Baseline celecoxib: 4.7 (1.0).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Abdominal pain, bloating, diarrhoea, dyspepsia, flatulence, gastric pain, gastrointestinal disorder, heartburn at 12 weeks; Group 1: 6/463, Group 2: 37/233; Comments: Ketoprofen 50mg: All = 3, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 0, flatulence = 0, gastric pain = 0, gastrointestinal disorder = 0, heartburn = 1. Ketoprofen 100mg: All = 3, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 0, flatulence = 0, gastric pain = 0, gastrointestinal disorder = 0, heartburn = 1. Celecoxib: All = 37, abdominal pain = 4, bloating = 2, diarrhoea = 3, dyspepsia = 5, flatulence = 0, gastric pain = 9, gastrointestinal disorder = 4, heartburn = 7.

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Vascular disorders at 12 weeks; Group 1: 1/463, Group 2: 4/233; Comments: Ketoprofen 50mg: 1. Ketoprofen 100mg: 0. Celecoxib: 4
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: All nervous system disorders (including headaches) at 12 weeks; Group 1: 0/463, Group 2: 2/233; Comments: Ketoprofen 50mg: All = 0, headache = 0. Ketoprofen 100mg: All = 0, headache = 0. Celecoxib: All = 2, headache = 0.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus PLACEBO (GEL)

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -1.9 (SD 1.67); n=463, Group 2: mean -1.8 (SD 1.67); n=472; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Baseline ketoprofen 50mg: 4.7 (1.1). Baseline ketoprofen 100mg: 4.8 (0.9). Baseline placebo 2.2g: 4.8 (1.0). Baseline placebo 4.4g: 4.9 (1.1).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 39, Reason: Placebo gel 2.2g: Administrative = 5, AE = 6, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1. Placebo gel 4.4g: Administrative = 7, AE = 7, Ineffective = 6, AE+Ineffective = 2, Lost to follow-up = 0.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Knee: Abdominal pain, bloating, diarrhoea, dyspepsia, flatulence, gastric pain, gastrointestinal disorder, heartburn at 12 weeks; Group 1: 6/463, Group 2: 9/472; Comments: Ketoprofen 50mg: All = 3, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 0, flatulence = 0, gastric pain = 0, gastrointestinal disorder = 0, heartburn = 1. Ketoprofen 100mg: All = 3, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 0, flatulence = 0, gastric pain = 0, gastrointestinal disorder = 0, heartburn = 1. Placebo gel 2.2g: All = 2, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 0, flatulence = 0, gastric pain = 1, gastrointestinal disorder = 0, heartburn = 0. Placebo gel 4.4g: All = 7, abdominal pain = 0, bloating = 0, diarrhoea = 0, dyspepsia = 1, flatulence = 1, gastric pain = 2, gastrointestinal disorder = 1, heartburn = 0.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 39, Reason: Placebo gel 2.2g: Administrative = 5, AE = 6, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1. Placebo gel 4.4g: Administrative = 7, AE = 7, Ineffective = 6, AE+Ineffective = 2, Lost to follow-up = 0.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Vascular disorders at 12 weeks; Group 1: 1/463, Group 2: 3/472; Comments: Ketoprofen 50mg: 1. Ketoprofen 100mg: 0. Placebo gel 2.2g: 2. Placebo gel 4.4g: 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 39, Reason: Placebo gel 2.2g: Administrative = 5, AE = 6, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1. Placebo gel 4.4g: Administrative = 7, AE = 7, Ineffective = 6, AE+Ineffective = 2, Lost to follow-up = 0.

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: All nervous system disorders (including headaches) at 12 weeks; Group 1: 0/463, Group 2: 4/472; Comments: Ketoprofen 50mg: All = 0, headache = 0. Ketoprofen 100mg: All = 0, headache = 0. Placebo gel 2.2g: All = 3, headache = 2. Placebo gel 4.4g: All = 1, headache = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 39, Reason: Ketoprofen 50mg: Administrative = 3, AE = 3, Ineffective = 7, AE+Ineffective = 0, Lost to follow-up = 1. Ketoprofen 100mg: Administrative = 7, AE = 13, Ineffective = 4, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 39, Reason: Placebo gel 2.2g: Administrative = 5, AE = 6, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1. Placebo gel 4.4g: Administrative = 7, AE = 7, Ineffective = 6, AE+Ineffective = 2, Lost to follow-up = 0.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO (ORAL)

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -1.9 (SD 1.62); n=233, Group 2: mean -1.42 (SD 1.62); n=227; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Baseline celecoxib: 4.7 (1.0). Baseline placebo: 4.8 (1.0).

Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 38, Reason: Placebo (oral): Administrative = 4, AE = 13, Ineffective = 19, AE+Ineffective = 2, Lost to follow-up = 0. Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Abdominal pain, bloating, diarrhoea, dyspepsia, flatulence, gastric pain, gastrointestinal disorder, heartburn at 12 weeks; Group 1: 37/233, Group 2: 33/227; Comments: Celecoxib: All = 37, abdominal pain = 4, bloating = 2, diarrhoea = 3, dyspepsia = 5, flatulence = 0, gastric pain = 9, gastrointestinal disorder = 4, heartburn = 7. Placebo (oral): All = 33, abdominal pain = 3, bloating = 3, diarrhoea = 3, dyspepsia = 3, flatulence = 5, gastric pain = 5, gastrointestinal disorder = 4, heartburn = 4. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 38, Reason: Placebo (oral): Administrative = 4, AE = 13, Ineffective = 19, AE+Ineffective = 2, Lost to follow-up = 0. Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Vascular disorders at 12 weeks; Group 1: 4/233, Group 2: 1/227; Comments: Celecoxib: 4. Placebo (oral): 1. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 38, Reason: Placebo (oral): Administrative = 4, AE = 13, Ineffective = 19, AE+Ineffective = 2, Lost to follow-up = 0. Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: All nervous system disorders (including headaches) at 12 weeks; Group 1: 2/233, Group 2: 5/227; Comments: Celecoxib: All = 2, headache = 0. Placebo: All = 5, headache = 5. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, baseline values for WOMAC scale and analgesic use; Group 1 Number missing: 25, Reason: Celecoxib: Administrative = 6, AE = 13, Ineffective = 5, AE+Ineffective = 0, Lost to follow-up = 1.; Group 2 Number missing: 38, Reason: Placebo (oral): Administrative = 4, AE = 13, Ineffective = 19, AE+Ineffective = 2, Lost to follow-up = 0.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Couto 2018[52]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	4 (4 studies pooled in the analysis) (n=818)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 7 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographic evidence (from the previous 3 years) was used to verify osteoarthritis of the hip or knee, including subchondral sclerosis, joint space narrowing, presence of osteophytes or marginal lipping, or cyst formation in the target joint typical of osteoarthritis stage I-III Kellgren and Lawrence changes
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were generally healthy and over the age of 25 years with confirmed osteoarthritis of the hip or knee. People had at least moderate pain on weight bearing, as well as episodic flare of osteoarthritis. People had not taken any analgesic medication within 72 hours of the baseline evaluation (except paracetamol), or had completed the required 2- or 4-day washout period before baseline.
Exclusion criteria	Stage IV osteoarthritis (Kellgren and Lawrence system); moderate-to-severe chronic low back pain; inflammatory joint diseases (including rheumatoid arthritis, gout, mixed connective tissue disease, seronegative spondyloarthropathy, psoriatic arthritis, or systemic lupus erythematosus); a recent traumatic injury; history of peptic ulceration within the previous 9 months; any GI surgery, complaints or dysfunction that could interfere with drug absorption; any other significant medical conditions; people with concurrent or recent history of severe illness (one that has not been stable for the last 3 months) especially if related to the hepatic, renal, cardiovascular, respiratory, haematopoietic or endocrine systems; people on a daily regimen of prescription NSAIDs for the past 3 months (occasional period use of NSAIDs "as needed" was allowed in two of the study, as was prophylactic use of acetylsalicylic acid 325mg/day); a history of hypersensitivity or intolerance to any of the drugs used in the trial, or to any other anti-inflammatory drug; people who had previously enrolled in these studies or had participated in another drug investigation of device trial within the preceding 4 weeks
Recruitment/selection of patients	Four separate multi-center trials

Age, gender and ethnicity	Age - Mean (SD): 60.7 (12.9). Gender (M:F): 234:584. Ethnicity: Asian = 6, Black = 46, Caucasian = 740, Hispanic = 21, Other = 5
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (208) or knee (610)).
Extra comments	Severity: Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=428) Intervention 1: NSAIDs - Naproxen. Naproxen sodium 220mg two to three times per day (two times per day if age ≥65 years, three times per day if age <65 years). Duration 1 week. Concurrent medication/care: Use of aspirin ≤325mg/day was permitted. Indirectness: No indirectness (n=419) Intervention 2: Placebo. Matching placebo twice or three times daily. Duration 1 week. Concurrent medication/care: Use of aspirin ≤325mg/day was permitted. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Bayer Consumer Health)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Digestive system adverse events at 1 week; Group 1: 57/409, Group 2: 47/409
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, location of osteoarthritis and baseline outcome values; Group 1 Number missing: 19, Reason: Reason not given; Group 2 Number missing: 10, Reason: Reason not given

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Cardiovascular system adverse events at 1 week; Group 1: 4/409, Group 2: 6/409
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, location of osteoarthritis and baseline outcome values; Group 1 Number missing: 19, Reason: Reason not given; Group 2 Number missing: 10, Reason: Reason not given

Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Urogenital system adverse events at 1 week; Group 1: 6/409, Group 2: 6/409
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, location of osteoarthritis and baseline outcome values; Group 1 Number missing: 19, Reason: Reason not given; Group 2 Number missing: 10, Reason: Reason not given Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Nervous system adverse events at 1 week; Group 1: 20/409, Group 2: 21/409 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race, location of osteoarthritis and baseline outcome values; Group 1 Number missing: 19, Reason: Reason not given; Group 2 Number missing: 10, Reason: Reason not given	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study (subsidiary papers)	Cryer 2011[53] (Hochberg 2011[92])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (pooled analysis of 2 RCTs) (n=1219)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic, clinically diagnosis osteoarthritis of the knee (American College of Rheumatology functional class rating of I-III)
Stratum	Knee:
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged at least 50 years with a 6-month history of symptomatic, clinically diagnosed osteoarthritis of the knee (American College of Rheumatology functional class rating I-III) who had been receiving a stable dose of NSAIDs, COX-2 selective inhibitors or other oral analgesic therapy for at least 6 weeks
Exclusion criteria	A history of hypersensitivity, allergic reaction, or intolerance to any PPI or an NSAID (including aspirin); a presence of any uncontrolled acute or chronic medical illness; any gastrointestinal disorder or surgery; a history of peptic ulcer disease within 6 months prior to screning; recent history (in the past 3 months) of alcohol or drug abuse.
Recruitment/selection of patients	The study includes a 7-14 day washout period (for people previously taking medication). In order to be included people had to show evidence of an osteoarthritis pain flare.
Age, gender and ethnicity	Age - Mean (range): 61.9 (49-90). Gender (M:F): 441:778. Ethnicity: White = 973, Black = 200, Other = 46
Further population details	1. Age: Mixed 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: ACR functional class I-III (median = II) Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=492) Intervention 1: NSAID and gastroprotection - NSAID and proton pump inhibitor. Naproxen 500mg and esomeprazole 20mg twice daily with a placebo capsule once daily. Duration 12 weeks. Concurrent medication/care: Excluded

	concomitant medications included any other NSAID (other than low dose aspirin) or gastroptective agent, parenteral steroids, lithium, glucosamine and/or chondroitin sulfate, and anticoagulants. Incidental use of rescue antacid (≤6 tablets per day) and supplemental use of rescue paracetamol (≤3g/day) were allowed during the study. Concomitant use of oral prednisone (≤7.5mg/day), low dose aspirin (≤325mg/day) and antiplatelet agents (non-concomitant with aspirin) were allowed. Indirectness: No indirectness (n=494) Intervention 2: NSAIDs - Celecoxib. Celecoxib 200mg capsule once daily and placebo tablets twice daily. Duration 12 weeks. Concurrent medication/care: Excluded concomitant medications included any other NSAID (other than low dose aspirin) or gastroptective agent, parenteral steroids, lithium, glucosamine and/or chondroitin sulfate, and anticoagulants. Incidental use of rescue antacid (≤6 tablets per day) and supplemental use of rescue paracetamol (≤3g/day) were allowed during the study. Concomitant use of oral prednisone (≤7.5mg/day), low dose aspirin (≤325mg/day) and antiplatelet agents (non-concomitant with aspirin) were allowed. Indirectness: No indirectness (n=248) Intervention 3: Placebo. Placebo tablets twice daily and placebo capsule once daily. Duration 12 weeks. Concurrent medication/care: Excluded concomitant medications included any other NSAID (other than low dose aspirin) or gastroptective agent, parenteral steroids, lithium, glucosamine and/or chondroitin sulfate, and anticoagulants. Incidental use of rescue antacid (≤6 tablets per day) and supplemental use of rescue paracetamol (≤3g/day) were allowed during the study. Concomitant use of oral prednisone (≤7.5mg/day), low dose aspirin (≤325mg/day) and antiplatelet agents (non-concomitant with aspirin) were allowed. Indirectness: No indirectness
Funding	Study funded by industry (These studies were sponsored and conducted by POZEN Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NSAID AND PROTON PUMP INHIBITOR versus CELECOXIB

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper GI adverse events (abdominal discomfort, abdominal pain, abdominal tenderness, duodenal ulcer, dyspepsia, nausea, gastritis, gastroesophageal reflux disease, GI haemorrhage, hyperchlorhydria, stomach discomfort, upper abdominal pain, vomiting) at 12 weeks; Group 1: 87/490, Group 2: 94/488; Comments: Including ... Naproxen/Esomeprazole: Diarrhoea = 27, dyspepsia = 41, nausea = 17, upper abdominal pain = 20, constipation = 17, vomiting = 8. Celecoxib: Diarrhoea = 14, dyspepsia = 52, nausea = 15, upper abdominal pain = 21, constipation = 10, vomiting = 6.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Outcome includes GI bleeding and non-bleeding events that we are unable to separate - the majority at GI non-bleeding events; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 2, Reason: No reason given; Group 2 Number missing: 6, Reason: No reason given

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Peripheral oedema at 12 weeks; Group 1: 15/490, Group 2: 6/488
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 2, Reason: No reason given; Group 2 Number missing: 6, Reason: No reason given

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NSAID AND PROTON PUMP INHIBITOR versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper GI adverse events (abdominal discomfort, abdominal pain, abdominal tenderness, duodenal ulcer, dyspepsia, nausea, gastritis, gastroesophageal reflux disease, GI haemorrhage, hyperchlorhydria, stomach discomfort, upper abdominal pain, vomiting) at 12 weeks; Group 1: 87/490, Group 2: 49/246; Comments: Including: Naproxen/Esomeprazole: Diarrhoea = 27, dyspepsia = 41, nausea = 17, upper abdominal pain = 20, constipation = 17, vomiting = 8. Placebo: Diarrhoea = 9, dyspepsia = 30, nausea = 9, upper abdominal pain = 8, constipation = 3, vomiting = 5.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Outcome includes GI bleeding and non-bleeding events that we are unable to separate - the majority at GI non-bleeding events; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 2, Reason: No reason given; Group 2 Number missing: 2, Reason: No reason given

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Peripheral oedema at 12 weeks; Group 1: 15/490, Group 2: 3/246
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 2, Reason: No reason given; Group 2 Number missing: 2, Reason: No reason given

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper GI adverse events (abdominal discomfort, abdominal pain, abdominal tenderness, duodenal ulcer, dyspepsia, nausea, gastritis, gastroesophageal reflux disease, GI haemorrhage, hyperchlorhydria, stomach discomfort, upper abdominal pain, vomiting) at 12 weeks; Group 1: 94/488, Group 2: 49/246; Comments: Including ... Celecoxib: Diarrhoea = 14, dyspepsia = 52, nausea = 15, upper abdominal pain = 21, constipation = 10,

vomiting = 6. Placebo: Diarrhoea = 9, dyspepsia = 30, nausea = 9, upper abdominal pain = 8, constipation = 3, vomiting = 5. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Outcome includes GI bleeding and non-bleeding events that we are unable to separate - the majority at GI non-bleeding events; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 6, Reason: No reason given; Group 2 Number missing: 2, Reason: No reason given Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Peripheral oedema at 12 weeks; Group 1: 6/488, Group 2: 3/246 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, BMI, smoking history, low dose aspirin use at randomisation and American College of Rheumatology functional class; Group 1 Number missing: 6, Reason: No reason given; Group 2 Number missing: 2, Reason: No reason given	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	**Dehghan 2020[54]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=145)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis in at least one knee with orthopedic diagnosis based on radiological criteria in knee image, having experienced pain for at least 2 weeks before treatment and having age above 45 years.
Stratum	Knee

Subgroup analysis within study	Not applicable
Inclusion criteria	Primary osteoarthritis in at least one knee with orthopedic diagnosis based on radiological criteria in knee image, having experienced pain for at least 2 weeks before treatment and having age above 45 years.
Exclusion criteria	People who had secondary osteoarthritis; active liver or kidney disease; peptic ulcer; diabetes; thyroid and parathyroid diseases; coagulation disorders; consumed anticoagulant drugs; had history of ischaemic or haemorrhagic stroke or deep vein thrombosis; allergy to any anti-inflammatory drug; alcohol abuse; drug abuse; acute trauma; skin diseases; infection or wounds at the site where the gel was applied; used corticosteroids of any type and other topical drugs at the site where the gel was applied; orally used other analgesics; and other effective compounds for the treatment of osteoarthritis up to 10 days before beginning of the study; had pregnancy; history of local fractures; deformities leading to osteoarthritis and articular diseases.
Recruitment/selection of patients	People referred to Imam Ali Clinic affiliated with Shahrekord University of Medical Sciences and private offices across Shahrekord diagnosed with knee osteoarthritis. People were selected by convenience sampling.
Age, gender and ethnicity	Age - Range: 45-75. Gender (M:F): 100:45. Ethnicity: Not stated/unclear
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated/unclear Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=49) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac 1% gel - one 60 gram tube. Gel was applied three times a day for 6 weeks applied over 3-5 minutes each time.. Duration 6 weeks. Concurrent medication/care: All people were given celecoxib 200mg capsules daily in addition to the topical gels. Indirectness: No indirectness (n=48) Intervention 2: Placebo. Placebo gel - one 60 gram tube. Gel was applied three times a day for 6 weeks applied over 3-5 minutes each time.. Duration 6 weeks. Concurrent medication/care: All people were given celecoxib 200mg capsules daily in addition to the topical gels. Indirectness: No indirectness (n=48) Intervention 3: NSAID gels (topical - local) - Other. H. helix herbal remedy gel. Gel was applied three times a day for 6 weeks applied over 3-5 minutes each time.. Duration 6 weeks. Concurrent medication/care: All people were given celecoxib 200mg capsules daily in addition to the topical gels. Indirectness: No indirectness Comments: This intervention is not included in the protocol so will not be included in the analysis.
Funding	Academic or government funding (This work was supported by the Deputy of Research and Technology of Shahrekord University of Medical Sciences (grant number: 2854).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Visual analogue scale at 6 weeks; Group 1: mean -1.1 (SD 1.06); n=49, Group 2: mean -0.85 (SD 0.79); n=48; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline diclofenac: 3.67 (0.68). Baseline placebo: 3.6 (0.73). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, marital status, age, education level, occupation and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function at 6 weeks; Group 1: mean -32.04 (SD 18.38); n=49, Group 2: mean -11.13 (SD 14.8); n=48; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline diclofenac: 58.47 (14.55). Baseline placebo: 54.4 (15.56). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, marital status, age, education level, occupation and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Delemos 2011[55]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1011)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiologically confirmed American College of Rheumatology Functional Class I-III osteoarthritis of the hip and/or knee.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults with radiologically confirmed American College of Rheumatology Functional Class I-III osteoarthritis of the hip and/or knee if they had moderate or severe pain, defined as pain intensity of ≥40mm on a 100-mm visual analogue scale (0=no pain, 100=extreme pain) in the most painful knee or hip joint that warranted treatment with COX-2 inhibitors, NSAIDs, paracetamol, or opioid analgesics for at least 75 of 90 days preceding the screening visit. In addition, people were required to be able to discontinue paracetamol, NSAIDs, COX-2 inhibitors, opioids, and other analgesics (except aspirin ≤325mg once daily for cardiovascular prophylaxis). Females were required to have a negative pregnancy test at baseline and be postmenopausal for at least 1 year, be biologically or surgically sterile (hysterectomy or tubal ligation) or (if of childbearing potential) practice abstinence or use a medically acceptable form of contraception during the study.
Exclusion criteria	Any current medical condition other than osteoarthritis that was not well controlled with treatment or any clinically significant condition that in the investigator's opinion precluded study participation; inflammatory arthritis, gout, pseudo-gout, or Paget disease that would interfere with the assessment of pain and other symptoms of osteoarthritis; chronic pain syndrome or fibromyalgia; any other clinically significant form of joint disease or prior joint replacement surgery at the index joint (ie, the knee or hip joint with the most severe pain); anticipated need for any surgical or other invasive procedure to be performed on the index joint during the course of the study; dysphagia or difficulty swallowing tablets or capsules; intractable nausea or vomiting; intolerance or contraindications to tramadol, celecoxib, other COX-2 inhibitors, NSAIDs, or sulfonamides; treatment with an anticoagulant; ileostomy or colostomy; gastrointestinal ulceration or bleeding within the prior 90 days; any corticosteroid therapy within the prior month; intra-articular corticosteroid therapy in the index joint

	within the prior 2 months; intra-articular viscosupplementation in a nonindex joint within the prior 2 months or in the index joint within the prior 4 months; liver function tests >2 times the upper limit of normal; creatinine >1.9mg/dL; history of seizures; use in the prior 14 days of a monoaminoxidase inhibitor, tricyclic antidepressant, other tricyclic compound, neuroleptic, selective serotonin reuptake inhibitor, serotonin/norepinephrine reuptake inhibitor, anorectic, bupropion, carbamazepine, or quinidine; a history of substance abuse within the prior 6 months; cancer within the prior 3 years (excluding squamous or basal cell carcinoma of the skin); active neoplastic disease; chronic respiratory insufficiency.
Recruitment/selection of patients	The study was conducted by 70 investigators across the United States between September 2002 and August 2003.
Age, gender and ethnicity	Age - Mean (SD): 59.92 (10.96). Gender (M:F): 369:632. Ethnicity: 820 people were white, 121 were black, <10% were Asian, Hispanic or other race/ethnicity.
Further population details	1. Age: <75 years (Includes people up to the age of 80. The majority are less than 75 years (determined by SD).). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Majority knee (around 74%) the rest hip).
Extra comments	Severity: Functional class I-III (mean class II) Duration of symptoms: 8.1 (7.9) years.
Indirectness of population	No indirectness
Interventions	(n=606) Intervention 1: Strong opioids (oral) - Tramadol. Oral tramadol. Three different doses were given in three groups (100mg, 200mg and 300mg). 100mg tablets were taken up to three times a day (with placebo tablets being taken if they were in the 100mg and 200mg group). All groups received 1 placebo capsule per day.. Duration 12 weeks. Concurrent medication/care: All people took 3 tablets and 1 capsule per day (with relevant placebo as required dependent on medication being taken.) No concomitant or rescue analgesic therapy was permitted during the study. Aspirin up to 325mg/day for cardiovascular prophylaxis was allowed as was paracetamol up to 2g/day for up to 3 consecutive days for reasons other than osteoarthritis or chronic pain if absolutely necessary.. Indirectness: No indirectness Comments: As per the protocol, different doses/formulations were merged to form one group. Therefore, all tramadol groups will be analysed together as one class. (n=203) Intervention 2: NSAIDs - Celecoxib. Oral celecoxib 200mg once a day (as one capsule). They also received three placebo tablets during the day.. Duration 12 weeks. Concurrent medication/care: All people took 3 tablets and 1 capsule per day

	(with relevant placebo as required dependent on medication being taken.) No concomitant or rescue analgesic therapy was permitted during the study. Aspirin up to 325mg/day for cardiovascular prophylaxis was allowed as was paracetamol up to 2g/day for up to 3 consecutive days for reasons other than osteoarthritis or chronic pain if absolutely necessary.. Indirectness: No indirectness (n=202) Intervention 3: Placebo. Three placebo tablets and one placebo capsule per day. Duration 12 weeks. Concurrent medication/care: All people took 3 tablets and 1 capsule per day (with relevant placebo as required dependent on medication being taken.) No concomitant or rescue analgesic therapy was permitted during the study. Aspirin up to 325mg/day for cardiovascular prophylaxis was allowed as was paracetamol up to 2g/day for up to 3 consecutive days for reasons other than osteoarthritis or chronic pain if absolutely necessary.. Indirectness: No indirectness
Funding	Study funded by industry (This work was supported by Biovail Corporation and Ortho-McNeil Janssen Scientific Affairs, LLC who provided staff for medical writing support.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus CELECOXIB

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical component summary at 12 weeks; Group 1: mean 3.1 (SD 8.5); n=599, Group 2: mean 5.2 (SD 8.5); n=202; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: 2.8 (0.6). Reported tramadol 200mg: 3.1 (0.6). Reported tramadol 300mg: 3.5 (0.6). Reported celecoxib: 5.2 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.
- Actual outcome for Other: SF-36 mental component summary at 12 weeks; Group 1: mean -0.5 (SD 8.5); n=599, Group 2: mean -0.1 (SD 8.5); n=202; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: -0.9 (0.6). Reported tramadol 200mg: -0.3 (0.6). Reported tramadol 300mg: -0.3 (0.6). Reported celecoxib: -0.1 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36

component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -96.9 (SD 127); n=599, Group 2: mean -130 (SD 127.9); n=202; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: -82.5 (8.9). Reported tramadol 200mg: -90.4 (8.9). Reported tramadol 300mg: 117.8 (8.9). Reported celecoxib: -130 (9). Baseline (SD) tramadol 100mg: 298.4 (101.3). Baseline (SD) tramadol 200mg: 302.9 (96.1). Baseline (SD) tramadol 300mg: 306.2 (107.3). Baseline (SD) celecoxib: 286.9 (96.1).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -300.1 (SD 412.2); n=599, Group 2: mean -429.2 (SD 416.4); n=202; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: -272.3 (29.0). Reported tramadol 200mg: -271.0 (29.1). Reported tramadol 300mg: 357.2 (29.0). Reported celecoxib: -429.2 (29.3). Baseline (SD) tramadol 100mg: 1034.0 (341.6). Baseline (SD) tramadol 200mg: 1045.1 (319.9). Baseline (SD) tramadol 300mg: 1023.6 (364.7). Baseline (SD) celecoxib: 991.1 (351.1).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical component summary at 12 weeks; Group 1: mean 3.1 (SD 8.5); n=599, Group 2: mean 3 (SD 8.5); n=200; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from

this. Reported tramadol 100mg: 2.8 (0.6). Reported tramadol 200mg: 3.1 (0.6). Reported tramadol 300mg: 3.5 (0.6). Reported placebo: 3.0 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.
- Actual outcome for Other: SF-36 mental component summary at 12 weeks; Group 1: mean -0.5 (SD 8.5); n=599, Group 2: mean -0.3 (SD 8.5); n=200; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: -0.9 (0.6). Reported tramadol 200mg: -0.3 (0.6). Reported tramadol 300mg: -0.3 (0.6). Reported placebo: -0.3 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -96.9 (SD 127); n=599, Group 2: mean -94.9 (SD 125.9); n=200; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported tramadol 100mg: -82.5 (8.9). Reported tramadol 200mg: -90.4 (8.9). Reported tramadol 300mg: 117.8 (8.9). Reported placebo: -94.9 (8.9). Baseline (SD) tramadol 100mg: 298.4 (101.3). Baseline (SD) tramadol 200mg: 302.9 (96.1). Baseline (SD) tramadol 300mg: 306.2 (107.3). Baseline (SD) placebo: 300.8 (103.5).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -300.1 (SD 412.2); n=599, Group 2: mean -290.1 (SD 411.5); n=200; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD

calculated from this. Reported tramadol 100mg: -272.3 (29.0). Reported tramadol 200mg: -271.0 (29.1). Reported tramadol 300mg: 357.2 (29.0). Reported placebo: -290.1 (29.1). Baseline (SD) tramadol 100mg: 1034.0 (341.6). Baseline (SD) tramadol 200mg: 1045.1 (319.9). Baseline (SD) tramadol 300mg: 1023.6 (364.7). Baseline (SD) placebo: 1019.0 (354.7).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 289, Reason: 7 people had no dose documented. The remaining 599 were included in the ITT analysis. However, 282 discontinued prematurely: 132 for adverse events, 106 for lack of efficacy, 12 for subject choice, 32 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical component summary at 12 weeks; Group 1: mean 5.2 (SD 8.5); n=202, Group 2: mean 3 (SD 8.5); n=200; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported celecoxib: 5.2 (0.6). Reported placebo: 3.0 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.
- Actual outcome for Other: SF-36 mental component summary at 12 weeks; Group 1: mean -0.1 (SD 8.5); n=202, Group 2: mean -0.3 (SD 8.5); n=200; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported celecoxib: -0.1 (0.6). Reported placebo: -0.3 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -130 (SD 127.9); n=202, Group 2: mean -94.9 (SD 125.9); n=200; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported

celecoxib: -130 (9). Reported placebo: -94.9 (8.9). Baseline (SD) celecoxib: 286.9 (96.1). Baseline (SD) placebo: 300.8 (103.5). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other. Protocol outcome 3: Physical function at ≤3- or >3- months - Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -429.2 (SD 416.4); n=202, Group 2: mean -290.1 (SD 411.5); n=200; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Reports least square mean difference and standard error. SD calculated from this. Reported celecoxib: -429.2 (29.3). Reported placebo: -290.1 (29.1). Baseline (SD) celecoxib: 991.1 (351.1). Baseline (SD) placebo: 1019.0 (354.7). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values for age, sex, race/ethnicity, index joint, osteoarthritis duration, functional class, WOMAC pain and physical function, and patient global assessment. Does not report SF-36 component scores at baseline.; Group 1 Number missing: 68, Reason: 1 person had no dose documented. The remaining 202 were included in the ITT analysis. However, 67 discontinued prematurely. 20 for adverse events, 30 for lack of efficacy, 2 for subject choice, 15 for other.; Group 2 Number missing: 99, Reason: 2 people had no dose documented. The remaining 200 were included in the ITT analysis. However, 97 discontinued prematurely. 15 for adverse events, 65 for lack of efficacy, 4 for subject choice, 13 for other.	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Dickson 1991[56]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=235)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Well documented, mild osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and females between 18 and 86 years of age with well documented, mild osteoarthritis of the knee.
Exclusion criteria	Use of NSAIDs (including aspirin) in the seven days prior to, and at anytime during the study; additional joint disease other than osteoarthritis; defects in the skin over the affected joints; known hypersensitivity to NSAIDs; a history of significant gastrointestinal disease including peptic ulceration; concurrent medication with anticoagulants or lithium; pregnancy or lactation.
Recruitment/selection of patients	All people were required to undergo a sever day washout period free of any anti-inflammatory medication prior to starting study drugs. Baseline assessments were made after this.
Age, gender and ethnicity	Age - Mean (SD): 62.5 (11.5). Gender (M:F): 80:155. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range, 21-86 years). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Mild Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=117) Intervention 1: NSAID gels (topical - local) - Other. Pirocixam gel - 1 gram applied to the affected knee three times daily and one placebo tablet three times daily. Duration 4 weeks. Concurrent medication/care: During the washout period and throughout the trial, paracetamol up to 4 grams per day was allowed as rescue analgesia. Indirectness: No indirectness

	(n=118) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 400mg three times daily with placebo gel three times daily. Duration 4 weeks. Concurrent medication/care: During the washout period and throughout the trial, paracetamol up to 4 grams per day was allowed as rescue analgesia. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a grant from Pfizer Ltd., Sandwich, Kent, United Kingdom)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PIROXICAM GEL versus IBUPROFEN

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper gastrointestinal and other gastrointestinal adverse events at 4 weeks; Group 1: 15/117, Group 2: 11/118; Comments: Piroxicam gel: Upper gastrointestinal = 12, other gastrointestinal = 3. Ibuprofen: Upper gastrointestinal = 10, other gastrointestinal = 1
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender and age; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Dependent oedema at 4 weeks; Group 1: 0/117, Group 2: 2/118
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender and age; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Central nervous system adverse events at 4 weeks; Group 1: 7/117, Group 2: 8/118
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender and age; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Dieppe 1993[57]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=89)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic knee osteoarthritis of more than 3 month duration. Received radiographs at the start of the study. Osteoarthritis was diagnosed on combined radiological and clinical grounds which included the presence of joint space narrowing and osteophytosis in at least one tibiofemoral compartment of the knee joint, and pain on joint use in all cases.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either sex, over the age of 25, with symptomatic knee osteoarthritis of more than 3 months duration and currently on regular NSAID treatment.
Exclusion criteria	Any contraindication to the use of diclofenac; any person who was unsuitable for the study; people who were currently being considered for knee surgery and who felt like it was unlikely they would be available over a 2 year period
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 63.1 (40-82). Gender (M:F): 24:65. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grades 1-4 (median grade 2-3) Duration of symptoms (mean): 9.8 years
Indirectness of population	No indirectness
Interventions	(n=45) Intervention 1: NSAIDs - Diclofenac. Diclofenac 100mg/day (slow release formulation). Duration 2 years. Concurrent medication/care: At entry, people stopped their previous NSAID tablet medication and switched immediately to the trial tablet. Tablets of 500mg paracetamol were provided as escape analgesia and people were instructed to take up to a maximum of eight tablets in 24 hours if necessary for pain relief.. Indirectness: No indirectness (n=44) Intervention 2: Placebo. Placebo once daily. Duration 2 years. Concurrent

	medication/care: At entry, people stopped their previous NSAID tablet medication and switched immediately to the trial tablet. Tablets of 500mg paracetamol were provided as escape analgesia and people were instructed to take up to a maximum of eight tablets in 24 hours if necessary for pain relief.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal side effects at 2 years; Group 1: 6/45, Group 2: 5/44; Comments: "Most withdrawals (as defined under assessment) occurred within the first 6 months the (two) main reason(s) being...side effects (five placebo, six active). The major side effects experienced were gastrointestinal."
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Only reports overall values for gender, age and mean disease duration; Group 1 Number missing: 14, Reason: 3 withdrew due to lack of efficacy, 6 due to side effects, 2 due to abnormal liver function tests, 3 due to lack of compliance; Group 2 Number missing: 24, Reason: 12 withdrew due to lack of efficacy, 5 due to side effects, 1 due to abnormal liver function tests, 3 due to lack of compliance, 3 due to intercurrent disease

Protocol outcome 2: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Abnormal liver function tests at 2 years; Group 1: 2/45, Group 2: 1/44
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Only reports overall values for gender, age and mean disease duration; Group 1 Number missing: 12, Reason: 3 withdrew due to lack of efficacy, 6 due to side effects, 3 due to lack of compliance; Group 2 Number missing: 23, Reason: 12 withdrew due to lack of efficacy, 5 due to side effects, 3 due to lack of compliance, 3 due to intercurrent disease

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Doherty 2011[58]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=892)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	40 years and older; knee pain for most of the past 3 months and on 4 of 7 preceding days; discontinuation of current analgesics; Steinbrocker functional capacity classification of I-III; pain affecting the index knee (after a washout period if currently taking analgesics) of 30mm or greater and 80mm or less on a 100mm visual analogue scale over the previous 48 hours for one or more of the following: walking on a flat surface, going up/down stairs, at night, sitting, lying, standing upright.
Exclusion criteria	Concomitant rheumatic, malignant, gastrointestinal, renal or hepatic conditions; taking anticoagulants (except ≤325mg aspirin/day); recent receipt of any drug to modify joint structure/function
Recruitment/selection of patients	Recruited from Nottinghamshire by postal questionnaire or local press/radio advertisements and the remainder by a site management organisation (Synexus Clinical Research, UK) by press/radio advertisements and general practitioner identification
Age, gender and ethnicity	Age - Mean (range): 60.6 (40-84). Gender (M:F): 455:437. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional classification I-III Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=224) Intervention 1: NSAIDs - Ibuprofen. Two ibuprofen 200mg capsules three times a day with water (≥6 hours between doses). Duration 13 weeks. Concurrent medication/care: Concomitant medication was allowed apart from anticoagulants (except ≤325mg aspirin/day). Indirectness: No indirectness

	(n=222) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol two 500mg capsules three times a day (≥6 hours between doses). Duration 13 weeks. Concurrent medication/care: Concomitant medication was allowed apart from anticoagulants (except ≤325mg aspirin/day). Indirectness: No indirectness (n=446) Intervention 3: NSAIDs - Other. Either 1 or 2 combination tables of paracetamol and ibuprofen three times a day (≥6 hours between doses). Duration 13 weeks. Concurrent medication/care: Concomitant medication was allowed apart from anticoagulants (except ≤325mg aspirin/day). Indirectness: No indirectness Comments: This intervention group is not included in the protocol and is reported for completeness
Funding	Study funded by industry (The trial was sponsored by Reckitt Benckiser Healthcare International Ltd)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -17.6 (SD 19.6); n=162, Group 2: mean -15.9 (SD 16.3); n=136; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 44.0 (15.2). Baseline paracetamol: 43.0 (14.9).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 62, Reason: Including (some participants may be excluded for multiple reasons): Adverse events = 34, insufficient pain relief = 17, withdrew consent = 8, lost to follow up = 4, protocol violation = 0, death/serious AE = 1, investigator decision = 3, other = 5; Group 2 Number missing: 85, Reason: Including (some participants may be excluded for multiple reasons): Adverse events = 39, insufficient pain relief = 29, withdrew consent = 13, lost to follow up = 5, protocol violation = 3, death/serious AE = 2, investigator decision = 3, other = 9

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean -13 (SD 17.1); n=158, Group 2: mean -12.7 (SD 17.2); n=133; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 42.8 (18.5). Baseline paracetamol: 42.7 (18.9).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 62, Reason: Including (some participants may be excluded for multiple reasons): Adverse events = 34, insufficient pain relief = 17, withdrew consent = 8, lost to follow up = 4, protocol violation = 0, death/serious AE = 1, investigator decision = 3, other = 5; Group 2 Number missing: 85, Reason: Including (some participants may be excluded for multiple reasons): Adverse events = 39, insufficient pain relief = 29, withdrew consent = 13, lost to follow up = 5, protocol violation = 3, death/serious AE = 2, investigator decision = 3, other = 9

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Dore 1995[59]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=254)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic, radiologically confirmed osteoarthritis of the knee that required NSAID treatment
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women 40 years or older if they have symptomatic, radiologically confirmed osteoarthritis of the knee that required NSAID treatment
Exclusion criteria	People with a connective tissue disorder; people scheduled for joint replacement within 1 year; a history of gastrointestinal bleeding, renal or hepatic impairment, drug-induced skin disorder, or any other illness likely to interfere with the evaluation or disposition of the study drug; had a contraindication to NSAIDs; had recent corticosteroid, investigational, anticoagulant, or cytoprotective therapy; women who were pregnant, breast-feeding, or of child-bearing potential and not using birth control
Recruitment/selection of patients	Previous arthritis medications were discontinued at the screening visit. Long-acting NSAIDs were stopped for at least 7 days; others were stopped for at least five half-lives of the arthritis medication. People had to have active disease after the washout period as defined by the following primary efficacy variables: overall assessment of condition by both the person and investigator had to be fair or poor; joint tenderness and walking pain had to be assessed as moderate to very severe.
Age, gender and ethnicity	Age - Other: Mean: 63.7. Gender (M:F): 94:160. Ethnicity: White = 225, Black = 19, Hispanic = 8, Other = 2
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=168) Intervention 1: NSAIDs - Naproxen. Naproxen 250mg two capsules twice daily or Etodolac 200mg two capsules twice daily. Duration 4 weeks. Concurrent

	medication/care: Concomitant NSAIDs (except for low-dose aspirin or an occasional non-anti-inflammatory analgesic for nonarthritic pain), corticosteroids, salicylate-containing topical preparations, and new or altered physiotherapy were prohibited. Maintenance of low-dose aspirin (maximum of 325mg/day) as antithrombosis prophylaxis and occasional use of non-anti-inflammatory analgesics for nonarthritic pain was permitted.. Indirectness: No indirectness Comments: Naproxen and etodolac groups were combined due to class effect as agreed in the protocol (n=86) Intervention 2: Placebo. Placebo two capsules twice daily. Duration 4 weeks. Concurrent medication/care: Concomitant NSAIDs (except for low-dose aspirin or an occasional non-anti-inflammatory analgesic for nonarthritic pain), corticosteroids, salicylate-containing topical preparations, and new or altered physiotherapy were prohibited. Maintenance of low-dose aspirin (maximum of 325mg/day) as antithrombosis prophylaxis and occasional use of non-anti-inflammatory analgesics for nonarthritic pain was permitted.. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a research grant from Wyeth-Ayerst Laboratories, Philadelphia, Pennsylvania)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/ETODOLAC versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Digestive system disturbances (including dyspepsia, flatulence and constipation) at 4 weeks; Group 1: 49/168, Group 2: 15/86; Comments: Etodolac: Digestive system disturbances total = 29, dyspepsia = 15, flatulence = 5, constipation = 2. Naproxen: Digestive system disturbances total = 20, dyspepsia = 9, flatulence = 1, constipation = 7. Placebo: Digestive system disturbances = 15, dyspepsia = 9, constipation = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, sex and race; Group 1 Number missing: 17, Reason: Insufficient therapeutic effect = 8, patient noncompliance = 3, admission criteria violation = 4, other = 2; Group 2 Number missing: 27, Reason: Insufficient therapeutic effect = 22, patient noncompliance = 1, admission criteria violation = 4

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nervous system disturbances at 4 weeks; Group 1: 10/168, Group 2: 5/86; Comments: Etodolac: Nervous system disturbances total = 4. Headache = 7. Naproxen: Nervous system disturbances total = 6. Headache = 6. Placebo: Nervous system disturbances total = 5. Headache = 11.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, sex and race; Group 1 Number missing: 17, Reason: Insufficient therapeutic effect = 8, patient noncompliance = 3, admission criteria violation = 4, other = 2; Group 2 Number

missing: 27, Reason: Insufficient therapeutic effect = 22, patient noncompliance = 1, admission criteria violation = 4	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	**Enteshari-moghaddam 2019[60]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=150)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate to severe idiopathic knee osteoarthritis able to walk with knee pain more than five by visual analogue scale or more than 48 score in the WOMAC score with radiographic evidence of osteoarthritis with Kellgren Lawrence score of III-IV.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Moderate to severe idiopathic knee osteoarthritis able to walk with knee pain more than five by visual analogue scale or more than 48 score in the WOMAC score with radiographic evidence of osteoarthritis with Kellgren Lawrence score of III-IV on x-ray.
Exclusion criteria	People with other rheumatologic diseases, including fibromyalgia, other knee or periarticular disease, lower limb fracture with knee joint involvement; osteoarthritis of the hip and ankle; history of knee surgery or trauma; radicular pain; intraarticular corticosteroid injection or physiotherapy in the previous three months; hyaluronic acid injection in the previous six months; psychologic disorders, including depression; balance control deficit; sensory and motor neurological deficits; cancer; people with underlying systemic disease, such as renal, hepatic or heart failure; uncontrolled blood pressure; diabetes mellitus; sever asthma in need of corticosteroid use; people who used corticosteroids in the last 6 weeks prior to the study; those with allergic reactions to any study medications.
Recruitment/selection of patients	People visiting rheumatology clinics of Ardabil University of Medical Sciences

Age, gender and ethnicity	Age - Mean (SD): 54.44 (7.17). Gender (M:F): 40:110. Ethnicity: Not stated/unclear
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade II-III Duration of symptoms (mean [SD]): 8.44 (4.07) years
Indirectness of population	No indirectness
Interventions	(n=50) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 30mg daily increased to a maximum of 60mg daily after two weeks (all people received the maximum dosage of each of the given medications after two weeks). . Duration 3 months. Concurrent medication/care: Opioid usage was not allowed during the intervention period. People were asked to continue their previous medications for osteoarthritis, except corticosteroids. People were asked to report the use of NSAIDs (ibuprofen with maximum dose of 2400mg daily, naproxen with maximum dose of 1500mg daily, indomethacin 150mg daily) during the study period.. Indirectness: No indirectness (n=50) Intervention 2: Anti-epileptics (oral) - Gabapentin. Gabapentin 300mg daily increased to a maximum of 600mg daily after two weeks (all people received the maximum dosage of each of the given medications after two weeks). . Duration 3 months. Concurrent medication/care: Opioid usage was not allowed during the intervention period. People were asked to continue their previous medications for osteoarthritis, except corticosteroids. People were asked to report the use of NSAIDs (ibuprofen with maximum dose of 2400mg daily, naproxen with maximum dose of 1500mg daily, indomethacin 150mg daily) during the study period.. Indirectness: No indirectness (n=50) Intervention 3: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg daily increased to a maximum of 2000mg daily after two weeks (all people received the maximum dosage of each of the given medications after two weeks). . Duration 3 months. Concurrent medication/care: Opioid usage was not allowed during the intervention period. People were asked to continue their previous medications for osteoarthritis, except corticosteroids. People were asked to report the use of NSAIDs (ibuprofen with maximum dose of 2400mg daily, naproxen with maximum dose of 1500mg daily, indomethacin 150mg daily) during the study period.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain percentage decrease at 3 months; Group 1: mean -78.29 (SD 10.06); n=50, Group 2: mean -50.32 (SD 10.78); n=50; WOMAC pain 0-100 Top=High is poor outcome; Comments: No baseline values reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical activity percentage decrease at 3 months; Group 1: mean -68.36 % (SD 11.69); n=50, Group 2: mean -58.82 % (SD 8.54); n=50; WOMAC physical activity 0-100 Top=High is poor outcome; Comments: No baseline values reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Fatigue and drowsiness at 3 months; Group 1: 7/50, Group 2: 0/50; Comments: SNRI: Fatigue = 2, drowsiness = 5.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GABAPENTIN versus SNRIS

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain percentage decrease at 3 months; Group 1: mean -73.94 % (SD 12.79); n=50, Group 2: mean -78.29 % (SD 10.06); n=50; WOMAC pain 0-100 Top=High is poor outcome; Comments: No baseline values reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical activity percentage decrease at 3 months; Group 1: mean -69.53 (SD 8.85); n=50, Group 2: mean -68.36 (SD 11.69); n=50; WOMAC physical activity 0-100 Top=High is poor outcome; Comments: No baseline values reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Fatigue and drowsiness at 3 months; Group 1: 4/50, Group 2: 7/50; Comments: Gabapentin: Fatigue = 2, drowsiness = 2. SNRI: Fatigue = 2, drowsiness = 5.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GABAPENTIN versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: WOMAC pain percentage decrease at 3 months; Group 1: mean -73.94 % (SD 12.79); n=50, Group 2: mean -50.32 % (SD 10.78); n=50; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline values not reported. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical activity percentage decrease at 3 months; Group 1: mean -69.53 % (SD 8.84); n=50, Group 2: mean -58.82 % (SD 8.54); n=50; WOMAC physical activity 0-100 Top=High is poor outcome; Comments: No baseline values reported. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Fatigue and drowsiness at 3 months; Group 1: 4/50, Group 2: 0/50; Comments: SNRI: Fatigue = 2, drowsiness = 2. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, disease duration, BMI and Kellgren Lawrence score.; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Esselinckx 1990[61]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=347)
Countries and setting	Conducted in Belgium; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with painful gonarthrosis or coxarthrosis, radiologically proved, necessitating almost continuous administration with an NSAID
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with painful gonarthrosis or coxarthrosis, radiologically proved, necessitating almost continuous administration with an NSAID
Exclusion criteria	People with unclear symptomatology; serious concomitant disease; evolutive gatroduodenal ulcers; pregnant or lactating women; people who had been given an intra-articular injection of a corticoid during the 6 weeks prior to the study
Recruitment/selection of patients	People were included if they were having almost continuous administration with an NSAID
Age, gender and ethnicity	Age - Mean (SD): 62.5 (10.1). Gender (M:F): 124:223. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range: 30-93). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee (around 67%) or hip (around 33%)).
Extra comments	Severity: Not stated Duration of symptoms: Median between 1 and 5 years (some less than 1 year, some more than 5 years).
Indirectness of population	No indirectness
Interventions	(n=258) Intervention 1: NSAIDs - Diclofenac. Either tenoxicam 20mg per day, tenoxicam 40mg per day or diclofenac retard 100mg per day. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: These were three groups that were combined in the analysis due to class effect as agreed in the protocol (n=89) Intervention 2: Placebo. Matching placebo. Duration 4 weeks. Concurrent

	medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC OR TENOXICAM versus PLACEBO Protocol outcome 1: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Oedema at 4 weeks; Group 1: 3/258, Group 2: 1/89; Comments: Tenoxicam 40mg: 2. Tenoxicam 20mg: 1. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, sex, type of osteoarthritis and duration of disease; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Essex 2012[63]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=322)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with osteoarthritis of the knee according to the American College of Rheumatology guidelines in a flare state, and with a functional capacity classification of I-III.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	African American participants aged at least 45 years, with osteoarthritis of the knee in a flare state, and with a functional capacity classification of I-III. For people receiving NSAID or analgesic therapy, a flare was demonstrate if the Physician's Global Assessment of Arthritis and Patient's Global Assessment of arthritis were both 'fair', 'poor' or 'very poor' at the baseline visit, and if the baseline Patient's Assessment of Arthritis pain visual analogue scale measurement was between 40 and 90mm, the Patient's Global assessment of Arthritis showed an increase of one or more grades and the Physician's Global Assessment of Arthritis showed an increase of one or more grades. For people who were not receiving treatment, a flare was defined if the Patient's Assessment of Arthritis Pain VAS was between 40 and 90mm and the Patient's and Physician's Global Assessment of Arthritis was 'poor' or 'very poor' and the Global Assessment of Arthritis was 'poor' or 'very poor'. People receiving NSAID or analgesic therapy were required to discontinue medication at least 48 hours before the baseline assessments.
Exclusion criteria	Functional capacity class IV; inflammatory arthritis (except those with gout or pseudo-gout who had not experienced an acute flare in the past 2 years, and those with regional pain syndrome or fibromyalgia); acute joint trauma with active symptoms within 3 months prior to the start of the study; oral corticosteroid therapy within 4 weeks, intramuscular or soft-tissue injections of corticosteroids within 2 months, intra-articular injections of corticosteroids within 3 months; intra-articular injections of hyaluronic acid in the index joint within 9 months before the first dose of study medication of paracetamol within 24 hours of the baseline visit; a diagnosis of oesophageal, gastric, pyloric channel or duodenal ulceration within 60 days before the first dose of study medication; unstable cardiovascular disease within 6 months before

	screening; history of gastrointestinal perforations, obstructions or bleeding; active GI, renal, hepatic or coagulation disorders; malignancy (unless surgically removed and no recurrence within 5 years); known sensitivity to NSAIDs, sulfonamides, aspirin or related compounds
Recruitment/selection of patients	For people receiving NSAID or analgesic therapy, a flare was demonstrate if the Physician's Global Assessment of Arthritis and Patient's Global Assessment of arthritis were both 'fair', 'poor' or 'very poor' at the baseline visit, and if the baseline Patient's Assessment of Arthritis pain visual analogue scale measurement was between 40 and 90mm, the Patient's Global assessment of Arthritis showed an increase of one or more grades and the Physician's Global Assessment of Arthritis showed an increase of one or more grades. For people who were not receiving treatment, a flare was defined if the Patient's Assessment of Arthritis Pain VAS was between 40 and 90mm and the Patient's and Physician's Global Assessment of Arthritis was 'poor' or 'very poor' and the Global Assessment of Arthritis was 'poor' or 'very poor'. People receiving NSAID or analgesic therapy were required to discontinue medication at least 48 hours before the baseline assessments.
Age, gender and ethnicity	Age - Mean (SD): 58.0 (8.5). Gender (M:F): 64:258. Ethnicity: All participants were African American
Further population details	1. Age: Mixed (Range = 45-83). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity classification of I-III Duration of symptoms (mean [SD]): 5.4 (5.8) years
Indirectness of population	No indirectness
Interventions	(n=255) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day with double dummy placebo regimen or Celecoxib 200mg once daily with double dummy placebo regimen. Duration 6 weeks. Concurrent medication/care: Paracetamol no more than 2g/day for no more than 3 consecutive days and not within 24 hours before any arthritis assessment was allowed during the study. People taking aspirin no more than 325mg/day were allowed to continue for the duration of the study if they had been taking a stable dose for at least 30 days before the first dose of study medication. Indirectness: No indirectness Comments: Naproxen and celecoxib groups were combined due to class effect as agreed in the protocol (n=67) Intervention 2: Placebo. Placebo twice a day. Duration 6 weeks. Concurrent

	medication/care: Paracetamol no more than 2g/day for no more than 3 consecutive days and not within 24 hours before any arthritis assessment was allowed during the study. People taking aspirin no more than 325mg/day were allowed to continue for the duration of the study if they had been taking a stable dose for at least 30 days before the first dose of study medication. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -5.3 (SD 4.5); n=249, Group 2: mean -4.7 (SD 4.8); n=65; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports least squares mean and standard error. Reported celecoxib (n=124): -4.9 (0.4). Reported naproxen (n=125): -5.7 (0.4). Reported placebo (n=65): -4.7 (0.6). Calculated SD celecoxib: 4.5. Calculated SD naproxen: 4.5. Calculated SD placebo: 4.8. Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report WOMAC subscales (only overall value). Reports age, gender, duration of osteoarthritis, global assessment results, functional capacity classification, VAS and total WOMAC score; Group 1 Number missing: 45, Reason: Naproxen: 7 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation). Celecoxib: 7 adverse events, 2 lack of efficacy, 0 abnormal laboratory data, 2 people defaulted, 14 other (includes protocol violation).; Group 2 Number missing: 23, Reason: Placebo: 2 adverse events, 10 lack of efficacy, 0 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation).

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -17.2 (SD 14.6); n=249, Group 2: mean -14.4 (SD 13.7); n=65; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports least squares mean and standard error. Reported celecoxib (n=124): -16.0 (1.3). Reported naproxen (n=125): -18.3 (1.3). Reported placebo (n=65): -14.4 (1.7). Calculated SD celecoxib: 14.5. Calculated SD naproxen: 14.5. Calculated SD placebo: 13.7. Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report WOMAC subscales (only overall value). Reports age, gender, duration of osteoarthritis, global assessment results, functional capacity classification, VAS and total WOMAC score; Group 1 Number missing: 46, Reason: Naproxen: 7 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation). Celecoxib: 7 adverse events, 2 lack of efficacy, 0 abnormal laboratory data, 2 people defaulted, 14 other (includes protocol violation).; Group 2 Number missing: 23, Reason: Placebo: 2 adverse events, 10 lack of efficacy, 0 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation).

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper gastrointestinal events (specifically moderate/severe nausea, moderate/severe abdominal pain, and/or moderate/severe

dyspepsia) at 6 weeks; Group 1: 6/255, Group 2: 4/67 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report WOMAC subscales (only overall value). Reports age, gender, duration of osteoarthritis, global assessment results, functional capacity classification, VAS and total WOMAC score; Group 1 Number missing: 46, Reason: Naproxen: 7 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation). Celecoxib: 7 adverse events, 2 lack of efficacy, 0 abnormal laboratory data, 2 people defaulted, 14 other (includes protocol violation).; Group 2 Number missing: 23, Reason: Placebo: 2 adverse events, 10 lack of efficacy, 0 abnormal laboratory data, 4 people defaulted, 7 other (includes protocol violation).	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Essex 2014[62]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=318)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee (diagnosed according to American College of Rheumatology criteria) who were determined to be in a flare state and had a functional capacity classification of I to III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged at least 45 years and of self-reported Hispanic descent with osteoarthritis of the knee (diagnosed according to American College of Rheumatology criteria) who were determined to be in a flare state and had a functional capacity classification of I to III. People actively being treated with an NSAID or other analgesic therapy discontinued treatment at least 48 hours prior to the baseline assessments. Eligible people indicated a Patient's Assessment of Arthritis pain visual analogue scale score between 40-90mm and had a minimum rating of 3 on the Physician's and Patient's Global Assessment of Arthritis at baseline.
Exclusion criteria	Inflammatory arthritis (except those with gout or pseudo-gout who had not experienced an acute flare in the past 2 years, and those with regional pain syndrome or fibromyalgia); acute joint trauma with active symptoms within 3 months prior to the start of the study; oral corticosteroid therapy within 4 weeks, intramuscular or soft-tissue injections of corticosteroids within 2 months, intraarticular injections of corticosteroids within 3 months, intraarticular injections of hyaluronic acid in the index joint within 9 months before the first dose of study medication or paracetamol within 24 hours of the baseline visit; a diagnosis of oesophageal, gastric, pyloric channel or duodenal ulceration within 60 days before the first dose of study medication; unstable cardiovascular disease within 6 months before screening; history of GI perforations, obstructions or bleeding; active GI, renal, hepatic or coagulation disorders; malignancy (unless surgically removed and no recurrence within 5 years); known sensitivity to NSAIDs, sulfonamides, aspirin or related compounds
Recruitment/selection of patients	People actively being treated with an NSAID or other analgesic therapy discontinued treatment at least 48 hours prior to the baseline assessments.

Age, gender and ethnicity	Age - Mean (SD): 60.4 (10.6). Gender (M:F): 107:211. Ethnicity: All participants were hispanic
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity classification I-III Duration of symptoms (mean [SD]): 6.0 (6.3) years.
Indirectness of population	No indirectness
Interventions	(n=256) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day with matching placebo or Celecoxib 200mg once a day with matching placebo. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The naproxen and celecoxib groups were combined due to class effect as agreed in the protocol (n=62) Intervention 2: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -4.6 (SD 4.5); n=254, Group 2: mean -4 (SD 4.7); n=61; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports least square mean and standard error. Calculated standard deviation from this. Reported celecoxib: -5.2 (0.4). Reported naproxen: -5.1 (0.4). Reported placebo: -4.0 (0.6). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, patient's and physician's global assessment, functional capacity classification, VAS score and WOMAC total score. Does not report subscale scores.; Group 1 Number missing: 66, Reason: Celecoxib: 3 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 11 subjects defaulted, 13 other. Naproxen: 9 adverse events, 3 lack of efficacy, 9 subject defaulted, 15 other.; Group 2 Number missing: 16, Reason: Placebo: 1 adverse events, 5 lack of efficacy, 5 subject defaulted, 5 other.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -16.15 (SD 15.8); n=254, Group 2: mean -11.1 (SD 14.8); n=61; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports least square mean and standard error. Calculated standard deviation from this. Reported celecoxib: -16.3 (1.4). Reported naproxen: -16.0 (1.4). Reported placebo: -11.1 (1.9). Baseline values not reported.

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, patient's and physician's global assessment, functional capacity classification, VAS score and WOMAC total score. Does not report subscale scores.; Group 1 Number missing: 66, Reason: Celecoxib: 3 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 11 subjects defaulted, 13 other. Naproxen: 9 adverse events, 3 lack of efficacy, 9 subject defaulted, 15 other.; Group 2 Number missing: 16, Reason: Placebo: 1 adverse events, 5 lack of efficacy, 5 subject defaulted, 5 other. Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Upper gastrointestinal events at 6 weeks; Group 1: 7/256, Group 2: 1/62; Comments: Naproxen: 4, celecoxib: 3, placebo: 1 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, patient's and physician's global assessment, functional capacity classification, VAS score and WOMAC total score. Does not report subscale scores.; Group 1 Number missing: 66, Reason: Celecoxib: 3 adverse events, 2 lack of efficacy, 1 abnormal laboratory data, 11 subjects defaulted, 13 other. Naproxen: 9 adverse events, 3 lack of efficacy, 9 subject defaulted, 15 other.; Group 2 Number missing: 16, Reason: Placebo: 1 adverse events, 5 lack of efficacy, 5 subject defaulted, 5 other.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Essex 2016[64]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=367)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Altman criteria for defining osteoarthritis of the knee according to the American Rheumatism Association
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Asian patients aged at least 45 years with defined criteria for knee osteoarthritis. All people were in an osteoarthritis flare state, and within a functional capacity classification of I-III (as classified by a physician).
Exclusion criteria	Not stated
Recruitment/selection of patients	All people were in an osteoarthritis flare state after a screening period where people were washed out of medication (for 1-14 days)
Age, gender and ethnicity	Age - Mean (SD): 64.8 (11.3). Gender (M:F): 120:247. Ethnicity: A study in Asian patients with knee osteoarthritis
Further population details	1. Age: Mixed (Based on range: 42-90). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of osteoarthritis (mean [SD]): 4.6 (4.7) years.
Indirectness of population	No indirectness
Interventions	(n=289) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice daily and corresponding placebo or Celecoxib 200mg once a day with corresponding placebo. Duration 6 weeks. Concurrent medication/care: Paracetamol (up to 2g/day) was permitted as a rescue analgesia for the treatment of arthritis symptoms during the pretreatment screening period. People were to discontinue use of paracetamol at least 24 hours prior to the baseline arthritis assessment. Indirectness: No indirectness (n=78) Intervention 2: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: Paracetamol (up to 2g/day) was permitted as a rescue analgesia for

	the treatment of arthritis symptoms during the pretreatment screening period. People were to discontinue use of paracetamol at least 24 hours prior to the baseline arthritis assessment. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Pfizer Inc.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Patient's assessment of arthritis pain (visual analogue scale) at 6 weeks; Group 1: mean -37.3 (SD 21.3); n=254, Group 2: mean -33.6 (SD 19.8); n=58; VAS 0-100 Top=High is poor outcome; Comments: Reports least square means and standard error. Calculated standard deviation from this. Reported naproxen: -37.5 (2.0). Reported celecoxib: -37.1 (2.0). Reported placebo: -33.6 (2.6). Calculated SD naproxen: 20.7. Calculated SD celecoxib: 22.0. Calculated SD placebo: 19.8. Baseline naproxen: 65.8 (11.7). Baseline celecoxib: 64.6 (12.2). Baseline placebo: 64.4 (13.0). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of osteoarthritis, patient's global assessment, physician's global assessment, functional capacity classification, VAS score and WOMAC total score; Group 1 Number missing: 66, Reason: Naproxen: 15 adverse events, 0 lack of efficacy, 7 subject defaulted, 17 others. Celecoxib: 10 adverse events, 0 lack of efficacy, 3 subject defaulted, 14 other.; Group 2 Number missing: 19, Reason: Placebo: 3 adverse events, 5 lack of efficacy, 4 subject defaulted, 7 other.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Famaey 1976[66]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=23)
Countries and setting	Conducted in Unknown; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks (2 weeks for each treatment)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip where the diagnosis was clinically and radiologically well established
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the hip where the diagnosis was clinically and radiologically well established and the disease was sufficiently painful to justify treatment with aspirin-like drugs
Exclusion criteria	People with serious hepatic, gastric, renal of haematological disorders
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Other: Mean: 66. Gender (M:F): 6:11. Ethnicity:
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: "6 were early cases and 17 were seriously affected by the disease" Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=23) Intervention 1: NSAIDs - Ketoprofen. Ketoprofen 50mg three times a day orally. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=23) Intervention 2: Placebo. Placebo three times a day. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN versus PLACEBO	

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Hip: Severe gastric pain and indigestion, nausea and vomiting, epigastric pain, regurgitation and indigestion at 2 weeks; Group 1: 7/20, Group 2: 2/20; Comments: Ketoprofen: 2 had severe gastric pain and indigestion, 3 had nausea (with 1 vomiting), 1 epigastric pain and 1 regurgitation and indigestion. Placebo: 1 indigestion, 1 mild gastric pain Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender and mean age for everyone in the study; Group 1 Number missing: 3, Reason: 3 excluded because they failed to finish the treatment; Group 2 Number missing: 3, Reason: 3 excluded because they failed to finish the treatment Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Hip: Hypertension with fluid retention at 2 weeks; Group 1: 1/20, Group 2: 0/20 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender and mean age for everyone in the study; Group 1 Number missing: 3, Reason: 3 excluded because they failed to finish the treatment; Group 2 Number missing: 3, Reason: 3 excluded because they failed to finish the treatment	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Fishman 2007[67] (Kean 2009[100])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=552)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks maintenance phase (6 day run in)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Based on the criteria of the American College of Rheumatology clinical classification
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged between 40 and 75 years with pain associated with osteoarthritis of the knee based on the criteria of the American College of Rheumatology criteria
Exclusion criteria	People with arthritic conditions other than osteoarthritis; people who had a history of seizures or who had taken any medication that reduces the seizure threshold within 3 weeks before randomisation; evidence of effusion >15cc on physical examination of the knee; a BMI ≥38; major illness requiring hospitalisation within 3 months before screening; unwillingness to discontinue pain medication or other medication being taken for the treatment of osteoarthritis; previous or current substance abuse or dependence other than nicotine; significant bowel, renal or liver disease; allergy or adverse reaction to opiates; any other condition that in the opinion of the investigator would have adversely affected the person's ability to complete the study
Recruitment/selection of patients	Recruited from 74 active centers in the United States
Age, gender and ethnicity	Age - Mean (SD): 61.2 (9.3). Gender (M:F): 206:333. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=325) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol Contramid OAD 100mg, 200mg and 300mg (once daily). During the 6 day run in, the dose was titrated by 100mg increments every 2 to 3 days until the randomised dose was reached. Duration 12 weeks. Concurrent medication/care: Rescue medication for pain due to

	osteoarthritis was not permitted. Short-acting analgesics for acute pain due to conditions other than osteoarthritis were allowed up to 3 days. However, use of the short-acting analgesics had to be stopped 3 days in advance of any study visit. Indirectness: No indirectness Comments: Consists of 3 groups which have been combined together for the analysis as agreed in the protocol (n=227) Intervention 2: Placebo. Matching placebo once daily. Duration 12 weeks. Concurrent medication/care: Rescue medication for pain due to osteoarthritis was not permitted. Short-acting analgesics for acute pain due to conditions other than osteoarthritis were allowed up to 3 days. However, use of the short-acting analgesics had to be stopped 3 days in advance of any study visit. Indirectness: No indirectness
Funding	Study funded by industry (This research was funded by Labopharm)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: WOMAC pain score at 12 weeks; Group 1: mean -43.5 (SD 45.7); n=315, Group 2: mean -32.3 (SD 48.2); n=224; WOMAC pain score 0-500 Top=High is poor outcome; Comments: Tramadol values combined. Tramadol 100mg: -41.6 (50.2). Tramadol 200mg: -42.8 (46.4). Tramadol 300mg: -46.0 (39.9). Baseline tramadol 100mg: 287.8 (78.8). Baseline tramadol 200mg: 283.8 (81.7). Baseline tramadol 300mg: 314.4 (97.1). Baseline placebo: 300.7 (88.8). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, and WOMAC pain score at baseline; Group 1 Number missing: 148, Reason: Adverse events = 68, treatment failure = 43, patient request = 22, investigator initiated = 15; Group 2 Number missing: 93, Reason: Adverse events = 17, treatment failure = 47, patient request = 9, investigator initiated = 19, death = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Fleischmann 1997[70]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=279)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographically confirmed symptomatic osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women 40 years of age or older with radiologically confirmed symptomatic osteoarthritis of the knees that required treatment with aspirin and/or other NSAIDs for at least 3 months prior to screening or people who clearly were in need of such medication but were not actually taking it if symptomatic for at least 3 months. People had to be able to perform usual activities of self-care but could have limits in vocational and avocational activity. The person's osteoarthritis could be primary or secondary to trauma with the causative trauma having to had occurred at least 5 years prior to entry to the study.
Exclusion criteria	Other forms of arthritis; concomitant disease that could affect the joints (e.g. psoriasis, syphilitic neuropathy, ochronosis, inflammatory bowel disease, acute trauma, or metabolic bone disease except osteoporosis); other connective tissue disorders; significant disease or major illness within the past 5 years; significant complications of diabetes; recent severe infections including tuberculosis; a history of major gastrointestinal bleeding or active peptic ulcer disease within the past 5 years or if they were scheduled for joint replacement within 1 year after enrollment; people who used insulin therapy or oral hypoglycaemic agents, oral or parenteral anticoagulant therapy, medications associated with drug-induced hepatitis, or cytoprotective agents or corticosteroids by any route except inhaled, nasal, or topical within 6 weeks prior to study entry; people who received an investigational drug within the past 90 days; women who were pregnant, nursing or of childbearing potential and who were not using a medically accepted means of birth control; people who were active alcoholics or drug abusers within the past year; people who had a significant psychiatric disorder; a history of a NSAID induced skin disorder; major surgery within 6 weeks prior to entry; conditions likely to interfere with drug pharmacokinetics or who required elective surgery during the study or new drug therapy upon entry into this study; people with

	serum creatinine levels above the laboratory's normal limits or alanine aminotransferase/aspartate aminotransferase levels greater than 1.5 times the upper limit of normal; people with a history of hypersensitivity reactions to aspirin or other NSAIDs manifested as asthma, laryngospasm, rhinitis or urticaria
Recruitment/selection of patients	Multicenter. Required people to have a washout of their previous NSAID medications before starting the trial. The washout took place for at least 2 days before the baseline visit (with 7 days for long acting NSAIDs). People who did not develop active disease after 2 weeks of washout were excluded. Active disease was defined by the following: the overall assessment of the person's condition by both the physician and patient had to be fair or poor, and joint tenderness and walking pain had to be assessed as moderate to very severe.
Age, gender and ethnicity	Age - Mean (SD): 62.3 (10.4). Gender (M:F): 87:192. Ethnicity: White = 225, Black = 34, Hispanic = 17, Other = 3
Further population details	1. Age: Mixed (34 people were over the age of 75 years). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 3 months..
Indirectness of population	No indirectness
Interventions	(n=185) Intervention 1: NSAIDs - Naproxen. Naproxen (Naprelan) 1000mg (two 500mg tablets once daily) with placebo (three tablets once daily) or Nabumetone 1500mg (three 500mg tablets once daily) with placebo (two tablets once daily). Duration 4 weeks. Concurrent medication/care: People were prohibited from taking any concomitant aspirin or aspirin containing products, analgesics for arthritis, or anti-inflammatory drugs other than their study medication. The only permitted exception was the established use of low-dose aspirin (maximum dose of 325mg) as antithrombotic prophylaxis and the occasional use of a pure analgesic at low doses for nonarthritic pain. No such analgesic was permitted within 12 hours of any study visit. The use of corticosteroids except by nasal or topical routes and ointments or creams containing salicylates or capsaicin was prohibited. Previously initiated physiotherapy and the use of established medications for nonarthritic conditions were allowed to continue. Histamine-2 receptor antagonists were permitted if people were already taking these agents and if they were given at a stable dosage throughout the study.. Indirectness: No indirectness Comments: Naproxen and nabumetone groups are reported separately but have been combined for this analysis as agreed in the protocol due to class effect

	(n=94) Intervention 2: Placebo. Placebo (five tablets once daily). Duration 4 weeks. Concurrent medication/care: People were prohibited from taking any concomitant aspirin or aspirin containing products, analgesics for arthritis, or anti-inflammatory drugs other than their study medication. The only permitted exception was the established use of low-dose aspirin (maximum dose of 325mg) as antithrombotic prophylaxis and the occasional use of a pure analgesic at low doses for nonarthritic pain. No such analgesic was permitted within 12 hours of any study visit. The use of corticosteroids except by nasal or topical routes and ointments or creams containing salicylates or capsaicin was prohibited. Previously initiated physiotherapy and the use of established medications for nonarthritic conditions were allowed to continue. Histamine-2 receptor antagonists were permitted if people were already taking these agents and if they were given at a stable dosage throughout the study.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by a grant from Wyeth-Ayerst Laboratories, Philadelphia, Pennsylvania)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN OR NABUMETONE versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Digestive system adverse events (including constipation, diarrhoea, dyspepsia, flatulence and nausea) at 4 weeks; Group 1: 34/185, Group 2: 18/94; Comments: Naproxen: Digestive system total = 18. Constipation = 3, diarrhoea = 5, dyspepsia = 2, flatulence = 3, nausea = 7. Nabumetone: Digestive system total = 16. Constipation = 0, diarrhoea = 7, dyspepsia = 5, flatulence = 1, nausea = 1. Placebo: Digestive system total = 17. Constipation = 2, diarrhoea = 7, dyspepsia = 0, flatulence = 1, nausea = 7.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, mean body weight, mean height and race; Group 1 Number missing: 25, Reason: Naproxen: Insufficient therapeutic effect = 5, lost to follow up = 1, unknown = 1. Nabumetone: Insufficient therapeutic effect = 13, patient noncompliance = 2, lost to follow up = 2, other = 1; Group 2 Number missing: 17, Reason: Insufficient therapeutic effect = 14, patient noncompliance = 1, lost to follow up = 1, unknown = 1

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nervous system adverse events at 4 weeks; Group 1: 5/185, Group 2: 5/94; Comments: Naproxen: Nervous system total = 3. Headache = 3. Nabumetone: Nervous system total = 2. Headache = 5. Placebo: Nervous system total = 5. Headache = 9.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, mean body weight, mean height and race; Group 1 Number missing: 25, Reason: Naproxen: Insufficient therapeutic effect = 5, lost to follow up = 1, unknown = 1. Nabumetone: Insufficient therapeutic effect = 13, patient noncompliance = 2, lost to follow up = 2, other = 1; Group 2 Number missing: 17, Reason: Insufficient therapeutic

effect = 14, patient noncompliance = 1, lost to follow up = 1, unknown = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Fleischmann 2001[69]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=129)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention time: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic (painful) osteoarthritis of the knee confirmed by the demonstration of osteophytes on knee radiographs taken within a year before enrollment
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 35 to 75 years with symptomatic (painful) osteoarthritis of the knee for ≥1 year if they had used NSAIDs for ≥3 months before study entry and were otherwise in good health. People were required to have at least moderate pain (pain intensity ≥2 on a scale of 0 to 4, with 0 being the least and 4 being the greatest pain intensity) in the target knee when their current analgesic was discontinued
Exclusion criteria	Any other form of arthritis; major trauma; infection; apparent avascular necrosis of the target knee within 6 months before study entry; anatomical deformities of the knee that could interfere with assessment; if they underwent arthroscopic procedures within 6 months or surgical procedures on the target knee within a year before the study; had knee replacements or were candidates for knee replacement within 1 year; people who received intra-articular injections of corticosteroids in the knee within 1 month; hyaluronic acid injections or systemic corticosteroids within 3 months; taking glucosamine within 10 days before the study; people who, in the opinion of the investigator, should not have been enrolled in the study based on the precautions, warnings or contraindications outlined in the tramadol package insert
Recruitment/selection of patients	Multicenter, outpatient, placebo controlled, parallel-group clinical trial
Age, gender and ethnicity	Age - Mean (SD): 62.5 (9.17). Gender (M:F): 49:80. Ethnicity: 117 White, 9 Black, 3 Other
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: 7.85 (6.58) years.

Indirectness of population	No indirectness
Interventions	(n=63) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol 50mg titrated up every 2 days to a target dose of 200mg/day (1 capsule 4 times a day). For the remaining 84 days during the double-blind phase the people were allowed to take 1-2 capsules 4 times a day (maximum 400mg/day).. Duration 13 weeks (and a 10 day titration period). Concurrent medication/care: No rescue medication was permitted. People were instructed to maintain a constant level of activity throughout the study. Physiotherapy (ie. hot/cold packs and massages) initiated before the double-blind phase was continued throughout the study, although it could not be initiated during the double blind phase. People were not to use other adjunctive therapy (eg. topical therapy, acupuncture) during the study. Indirectness: No indirectness (n=66) Intervention 2: Placebo. Matching placebo titrated similarly to tramadol. Duration 13 weeks (and a 10 day titration period). Concurrent medication/care: No rescue medication was permitted. People were instructed to maintain a constant level of activity throughout the study. Physiotherapy (ie. hot/cold packs and massages) initiated before the double-blind phase was continued throughout the study, although it could not be initiated during the double blind phase. People were not to use other adjunctive therapy (eg. topical therapy, acupuncture) during the study. Indirectness: No indirectness
Funding	Study funded by industry (Funding for this study was provided by a research grant from Ortho-McNeil Pharmaceutical, Raritan, New Jersey (Study #CAPSS-051).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean 3.92 (SD 2.12); n=63, Group 2: mean 4.89 (SD 2.34); n=66; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Baseline values not reported (paper states they are similar to each other)
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, duration of disease and pain intensity score after washout period. Does not report baseline WOMAC subscale scores.; Group 1 Number missing: 43, Reason: Reasons for discontinuing during the first 20 days (23 people) were given: 13 due to drug ineffective, 10 for adverse events, 5 for patient choice, 3 for other. The numbers do not add up.; Group 2 Number missing: 49, Reason: Reasons for discontinuing during the first 20 days (28 people) were given: 27 due to drug ineffective, 2 for adverse events, 6 for patient choice, 1 for protocol violation, 1 for other. The numbers do not add up.

Protocol outcome 2: Physical function at ≤3- or >3- months

- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean 4.19 (SD 2.06); n=63, Group 2: mean 4.92 (SD 2.29); n=66; WOMAC physical function subscale 0-10 Top=High is poor outcome; Comments: Baseline values not reported. The paper states that they are similar between groups. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, duration of disease and pain intensity score after washout period. Does not report baseline WOMAC subscale scores.; Group 1 Number missing: 43, Reason: Reasons for discontinuing during the first 20 days (23 people) were given: 13 due to drug ineffective, 10 for adverse events, 5 for patient choice, 3 for other. The numbers do not add up.; Group 2 Number missing: 49, Reason: Reasons for discontinuing during the first 20 days (28 people) were given: 27 due to drug ineffective, 2 for adverse events, 6 for patient choice, 1 for protocol violation, 1 for other. The numbers do not add up.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Fleischmann 2006[68]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1600)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with primary osteoarthritis of the knee (as confirmed by the American College of Rheumatology criteria)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged more than 18 years with a diagnosis of primary osteoarthritis of the knee (as confirmed by the American College of Rheumatology criteria) with symptoms for at least 3 months. Females had to be non-pregnant, non-lactating, and practicing a reliable method of contraception, or surgically sterile or postmenopausal for at least 12 months.
Exclusion criteria	People with secondary osteoarthritis; other types of disease in the target knee; primary fibromyaglai; clinically significant medical problems
Recruitment/selection of patients	People meeting the initial inclusion criteria underwent a screening period for 3-7 days during which NSAID and/or other analgesic therapy was not permitted, with the exception of rescue medication (paracetamol). At the end of the screening period, people with pain intensity (during the last 24 hours) in the target knee greater than or equal to 40mm on a 100-mm visual analogue scale were eligible for entry into the treatment phase
Age, gender and ethnicity	Age - Mean (SD): 61.1 (11.3). Gender (M:F): 539:1061. Ethnicity: White/Caucasian = 1233, Hispanic = 247, Black/African American = 101, Other = 19
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: 6.4 (7.4) years
Indirectness of population	No indirectness
Interventions	(n=446) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 13 weeks. Concurrent medication/care: Paracetamol (≤2 gram/day) was permitted as rescue medication during the screening and treatment periods. Concomitant therapy

	with any other NSAIDs (with the exception of low dose aspirin ≤325 mg/day for a cardiovascular indication) was prohibited. Other prohibited concomitant therapies included histamine-2 receptor blockers, proton pump inhibitors, misoprostol, sucralphate (antacids were permitted up to twice a week), intra-articular/peri-articular injections and corticosteroids (ocular, topical, nasal or inhaled corticosteroids were allowed if the dosage was stable at study entry and throughout the study period).. Indirectness: No indirectness (n=930) Intervention 2: NSAIDs - Other. Lumiracoxib 200mg once a day or 400mg once a day. Duration 13 weeks. Concurrent medication/care: Paracetamol (≤2 gram/day) was permitted as rescue medication during the screening and treatment periods. Concomitant therapy with any other NSAIDs (with the exception of low dose aspirin ≤325 mg/day for a cardiovascular indication) was prohibited. Other prohibited concomitant therapies included histamine-2 receptor blockers, proton pump inhibitors, misoprostol, sucralphate (antacids were permitted up to twice a week), intra-articular/peri-articular injections and corticosteroids (ocular, topical, nasal or inhaled corticosteroids were allowed if the dosage was stable at study entry and throughout the study period).. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK, therefore was not included in the analysis. However, it was reported for completeness. (n=232) Intervention 3: Placebo. Placebo once a day. Duration 13 weeks. Concurrent medication/care: Paracetamol (≤2 gram/day) was permitted as rescue medication during the screening and treatment periods. Concomitant therapy with any other NSAIDs (with the exception of low dose aspirin ≤325 mg/day for a cardiovascular indication) was prohibited. Other prohibited concomitant therapies included histamine-2 receptor blockers, proton pump inhibitors, misoprostol, sucralphate (antacids were permitted up to twice a week), intra-articular/peri-articular injections and corticosteroids (ocular, topical, nasal or inhaled corticosteroids were allowed if the dosage was stable at study entry and throughout the study period).. Indirectness: No indirectness
Funding	Study funded by industry (Study supported by Novartis Pharma AG)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -3.5 (SD 4.11); n=444, Group 2: mean -2.3 (SD 3.9); n=231; WOMAC pain

subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 10.3 (3.34). Baseline placebo: 9.9 (3.30). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BmI, disease duration, and baseline values of outcomes; Group 1 Number missing: 99, Reason: 99 withdrawn. 33 unsatisfactory therapeutic effect, 3 lost to follow up, 33 adverse events, 30 other; Group 2 Number missing: 68, Reason: 68 withdrawn, 35 unsatisfactory therapeutic effect, 3 lost to follow up, 17 adverse events, 13 others Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC difficulty in performing daily activities subscale at 13 weeks; Group 1: mean -11 (SD 13.06); n=444, Group 2: mean -6.1 (SD 11.71); n=231; WOMAC Difficulty in Performing Daily Activities subscale 0-100 Top=High is poor outcome; Comments: Baseline Celecoxib: 33.0 (8.65). Baseline placebo: 33.0 (8.51). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BmI, disease duration, and baseline values of outcomes; Group 1 Number missing: 99, Reason: 99 withdrawn. 33 unsatisfactory therapeutic effect, 3 lost to follow up, 33 adverse events, 30 other; Group 2 Number missing: 68, Reason: 68 withdrawn, 35 unsatisfactory therapeutic effect, 3 lost to follow up, 17 adverse events, 13 others Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Number of people with gastrointestinal adverse events at 13 weeks; Group 1: 96/444, Group 2: 46/231 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BmI, disease duration, and baseline values of outcomes; Group 1 Number missing: 99, Reason: 99 withdrawn. 33 unsatisfactory therapeutic effect, 3 lost to follow up, 33 adverse events, 30 other; Group 2 Number missing: 68, Reason: 68 withdrawn, 35 unsatisfactory therapeutic effect, 3 lost to follow up, 17 adverse events, 13 others	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Frakes 2011[72]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=524)
Countries and setting	Conducted in Puerto Rico, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who met the American College of Rheumatology clinical and radiographic criteria for osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women older than age 40 years who met the American College of Rheumatology clinical and radiographic criteria for osteoarthritis of the knee. People had knee pain for at least 14 days/month in the 3 months preceding study entry and reported use of oral NSAIDs for treatment of knee pain on most days during that time period.
Exclusion criteria	Those with a history of gout, pseudogout, inflammatory arthritis, end stage bone-on-bone osteoarthritis, knee surgery or intra-articular injection in the past 6 months; people with another chronic painful condition such as fibromyalgia that could interfere with assessment of the knee; those with a body mass index >40kg/m²; those who were nonambulatory or required equipment other than a single cane to walk; people using opioid analgesics for more than 3 days/week who were unwilling to taper and discontinue their use prior to randomisation; people with other serious or unstable conditions, such as those taking warfarin and those with anaemia at baseline (haemoglobin <110g/L for males and <100g/L for females), or history of peptic ulcer disease, bleeding disorder or other risk factor for bleeding.
Recruitment/selection of patients	People were recruited from 42 study sites in the US and Puerto Rico between November 2009 and April 2011
Age, gender and ethnicity	Age - Mean (SD): 61.0 (9.2). Gender (M:F): 225:299. Ethnicity: White = 424, African American = 84, Other = 15
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 9.5 (8.9) years. Trial registration: NCT01018680. Results are supported by those published on clinicaltrials.gov (https://clinicaltrials.gov/ct2/show/results/NCT01018680?term=NCT01018680&draw=2&rank=1).
Indirectness of population	No indirectness
Interventions	(n=264) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine - starting at 30mg/day for 1 week, followed by 60mg/day for 2 weeks and increased to 120mg/day if their mean average pain severity rating was at least 4

	during the previous week at week 3.. Duration 10 weeks. Concurrent medication/care: Before the study started, each person had their NSAID therapy optimised. During the study they stayed on the same dose throughout the study period. People continued omeprazole 20mg daily to reduce the risk of upper GI bleeding. Paracetamol (650mg orally every 6 hours, not to exceed 2.6g/day or to be used for more than 25 days total throughout the course of the study) was permitted as rescue medication. Routine use of other analgesic agents (including anticonvulsants, opioids and topical agents), muscle relaxants, and sedative hypnotics were not permitted. For pain related to acute trauma or minor surgery, opioids could be used for up to three consecutive days, with total use not to exceed 10 days. Continued use of routine medication was permitted, as was the use of herbal therapies and nonpharmacological treatments such as physical therapy if they had been used routinely prior to study entry.. Indirectness: No indirectness (n=260) Intervention 2: Placebo. Matching placebo once per day. Duration 10 weeks. Concurrent medication/care: Before the study started, each person had their NSAID therapy optimised. During the study they stayed on the same dose throughout the study period. People continued omeprazole 20mg daily to reduce the risk of upper GI bleeding. Paracetamol (650mg orally every 6 hours, not to exceed 2.6g/day or to be used for more than 25 days total throughout the course of the study) was permitted as rescue medication. Routine use of other analgesic agents (including anticonvulsants, opioids and topical agents), muscle relaxants, and sedative hypnotics were not permitted. For pain related to acute trauma or minor surgery, opioids could be used for up to three consecutive days, with total use not to exceed 10 days. Continued use of routine medication was permitted, as was the use of herbal therapies and nonpharmacological treatments such as physical therapy if they had been used routinely prior to study entry.. Indirectness: No indirectness
Funding	Study funded by industry (Financial support was provided by Eli Lilly and Company, Indianapolis, IN, USA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 10 weeks; Group 1: mean -22.05 (SD 17.7); n=258, Group 2: mean -15.6 (SD 14.4); n=256; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported SNRI: -22.05 (1.1). Reported placebo: -15.6 (0.9). Scores were normalised to a 0-100 scale. No baseline values.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1

Protocol outcome 2: Physical function at ≤3- or >3- months

- Actual outcome for Knee: WOMAC physical function subscale at 10 weeks; Group 1: mean -21.1 (SD 17.9); n=251, Group 2: mean -13.81 (SD 18); n=253; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports mean change scores and SE. Calculated SD from this. Reported SNRI: -25.1 (1.13). Reported placebo: -13.81 (1.13). Baseline SNRI: 37.4 (10.1). Baseline placebo: 37.5 (9.4). Scores were normalised to a 0-100 scale.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Pancreatitis (and cholecystitis) at 10 weeks; Group 1: 0/264, Group 2: 1/260
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Coronary artery disease, syncope at 10 weeks; Group 1: 2/264, Group 2: 0/260
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1

Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Urosepsis at 10 weeks; Group 1: 1/264, Group 2: 0/260
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1

Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

- Actual outcome for Knee: Fall, metastasis to central nervous system, cerebrovascular accident at 10 weeks; Group 1: 2/264, Group 2: 1/260; Comments: SNRIs: 1 fall, 1 metastasis to central nervous system. Placebo: 1 cerebrovascular accident. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, duration of osteoarthritis pain, body mass index, weekly average pain, WOMAC physical function score and medication usage. Does not report baseline values for WOMAC pain.; Group 1 Number missing: 75, Reason: Adverse events = 40, subject decision = 10, protocol violation = 8, sponsor decision = 7, entry criteria not met = 4, lost to follow-up = 3, lack of efficacy = 2, physician decision = 1; Group 2 Number missing: 61, Reason: Adverse event = 23, subject decision = 10, lack of efficacy = 8, sponsor decision = 6, lost to follow-up = 5, protocol violation = 5, physician decision = 3, entry criteria not met = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study (subsidiary papers)	Frestedt 2008[74] (Frestedt 2009[75])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=70)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee according to their previous medical history and the modified clinical criteria of the American College of Rheumatology
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either gender who voluntarily gave informed consent, were ambulatory, 25-75 years old, with normal digestion and absorption, diagnosed with moderate to severe osteoarthritis of the knee according to their previous medical history and the modified clinical criteria of the of the American College of Rheumatology and had a WOMAC score of no more than 75 in the target knee. People were asked to consume a diet with around 600mg calcium which was estimated to be 40-60% of their RDI per day
Exclusion criteria	Rheumatoid arthritis; gout; pseudogout; Paget's disease; seizure disorder; insulin dependent diabetes mellitus; uncontrolled hypertension; unstable cardiovascular disease; active hepatic or renal disease; active cancer and/or HIV infection; if they require prescription drugs to control pain; had other clinical trial or experimental treatments in the past 3 months; were pregnant, lactating, or at risk of becoming pregnant; if they had received NSAIDs within 48 hours, intramuscular/systemic corticosteroid injection within 4 weeks, or intraarticular hyaluronic acid injection within 4 months prior to enrollment
Recruitment/selection of patients	No specific selection criteria. No statement about glucosamine quality.
Age, gender and ethnicity	Age - Mean (SD): 59.2 (9.7). Gender (M:F): 33:37. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=19) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulfate (167mg D-glucosamine sulfate potassium salt + 267mg maltodextrin) - three capsules three times per day (total dose = 1500mg).. Duration 12 weeks. Concurrent medication/care: The rescue medication was paracetamol, 325mg, 1-2 tablets every 4-6 hours as needed for intractable pain. Indirectness: No indirectness (n=35) Intervention 2: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine and Aquamin, or Aquamin only, three capsules three times a day. Duration 12 weeks. Concurrent medication/care: The rescue medication was paracetamol, 325mg, 1-2 tablets every 4-6 hours as needed for intractable pain. Indirectness: No indirectness Comments: Glucosamine sulfate and Aquamin, and Aquamin only are not included in the intervention section of the protocol and so are not included in the analysis (n=16) Intervention 3: Placebo. Three placebo capsules three times a day. Duration 12 weeks. Concurrent medication/care: The rescue medication was paracetamol, 325mg, 1-2 tablets every 4-6 hours as needed for intractable pain. Indirectness: No indirectness
Funding	Study funded by industry (Marigot Ltd. provided funding for this clinical trial and the article processing charges to publish this work.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -12.6 (SD 16.3); n=19, Group 2: mean -2.9 (SD 19.9); n=16; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 60.6 (14.8). Baseline placebo: 50.0 (22.9).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 didn't finish the study. Reasons not given.; Group 2 Number missing: 7, Reason: 7 didn't finish the study. Reasons not given.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC activity subscale at 12 weeks; Group 1: mean -10.5 (SD 24); n=19, Group 2: mean -7 (SD 18.4); n=16; WOMAC activity subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 60.1 (13.9). Baseline placebo: 49.4 (23.1).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 didn't finish the study. Reasons not given.; Group 2 Number missing: 7, Reason: 7 didn't finish the study. Reasons not given. Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal adverse events at 12 weeks; Group 1: 0/19, Group 2: 0/16 Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 didn't finish the study. Reasons not given.; Group 2 Number missing: 7, Reason: 7 didn't finish the study. Reasons not given. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiovascular/hypertension adverse events at 12 weeks; Group 1: 0/19, Group 2: 0/16 Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 didn't finish the study. Reasons not given.; Group 2 Number missing: 7, Reason: 7 didn't finish the study. Reasons not given. Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Neurological adverse events at 12 weeks; Group 1: 1/19, Group 2: 1/16 Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 didn't finish the study. Reasons not given.; Group 2 Number missing: 7, Reason: 7 didn't finish the study. Reasons not given.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Friedmann 2011[76]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=412)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (after a 2 week open label period)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip and/or knee demonstrated by clinical and radiographic evidence according to the American College of Rheumatology criteria for diagnosis
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Male or nonpregnant, nonbreastfeeding females (40-75 years) who had experienced moderate to severe chronic pain for ≥3 months owing to osteoarthritis of the hip and/or knee as demonstrated by clinical and radiographic evidence according to the American College of Rheumatology criteria for diagnosis and who were regularly (≥4 days/week every week for 4 weeks prior to the screening visit) taking one or more of the following oral medications: nonsteroidal antiinflammatory drugs, cyclooxygenase-2 inhibitors, tramadol and opioids. People could not have received opioids within 72 hours of the screening visit.
Exclusion criteria	People who had a daily opioid dose equivalent to oxycodone >80mg for ≥4 days/week during the week before the initial screening visit; those who had an intra-articular injection of hyaluronic acid in the hip or knee in the 6 months prior to the screening visit, or corticosteroid therapy (oral, intra-articular, or parenteral) or an epidural or intrathecal infusion of any analgesic medications in the 1 month prior to the screening visit; and those with a positive urine drug screen for opiates, amphetamines, cocaine, cannabinoids, phenylcyclohexylpiperidine, or methadone at the baseline visit; high doses of sedatives, hypnotics, tranquilizers, phenothiazines and other agents compromising vasomotor tone
Recruitment/selection of patients	Used an enriched enrollment randomised design (had an initial open label phase where everyone was on the drug and was slowly uptitrated and they excluded people who didn't tolerate it at all during that phase).
Age, gender and ethnicity	Age - Mean (SD): 58.3 (8.2). Gender (M:F): 124:288. Ethnicity: White = 338, Black = 69, Native American/Alaskan native = 5

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Moderate to severe Duration of symptoms: Not stated explicitly. At least 3 months
Indirectness of population	No indirectness
Interventions	(n=205) Intervention 1: Strong opioids (oral) - Oxycodone. Oxycodone extended release - initially 5mg twice daily increased up to 20mg twice daily by increases of 5mg over 3-4 day intervals in 2 weeks.. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (≤325mg per day) was allowed for cardiovascular prophylaxis. A stable dose for more than 4 weeks was required for monoaminoxidase inhibitors, tricyclic antidepressants, serotonin reuptake inhibitors or other antidepressants, gabapentin, pregabalin, and glucosamine/chondroitin. Indirectness: No indirectness (n=207) Intervention 2: Placebo. Matching placebo twice daily. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (≤325mg per day) was allowed for cardiovascular prophylaxis. A stable dose for more than 4 weeks was required for monoaminoxidase inhibitors, tricyclic antidepressants, serotonin reuptake inhibitors or other antidepressants, gabapentin, pregabalin, and glucosamine/chondroitin. Indirectness: No indirectness
Funding	Study funded by industry (Funded by King Pharmaceuticals, which was acquired by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYCODONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain intensity (visual analogue scale) at 12 weeks; Group 1: mean -0.7 (SD 2.05); n=203, Group 2: mean -0.3 (SD 2.48); n=207; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline oxycodone = 5.2 (2.19). Baseline placebo = 5.4 (2.11).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, target osteoarthritis pain joint, pain intensity score, SF-12 health survery score, WOMAC scores; Group 1 Number missing: 70, Reason: Adverse events = 43, inadequate pain relief = 12, patient request = 8, protocol violation = 7; Group 2 Number missing: 75, Reason: Inadequate pain relief = 38, adverse events = 22, protocol violation = 6, other = 5, patient request = 4

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	GAIT trial: Clegg 2006[50] (Hochberg 2008[91], Sawitzke 2008[165], Sawitzke 2010[166])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1583)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical evidence (knee pain for at least 6 months and on the majority of days during the preceding month) and radiographic evidence (tibiofemoral osteophytes of at least 1 mm [Kellgren Lawrence grade 2-3]) osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People at least 40 years of age with clinical evidence (knee pain for at least 6 months and on the majority of days during the preceding month) and radiographic evidence (tibiofemoral osteophytes of at least 1 mm [Kellgren Lawrence grade 2-3]) osteoarthritis of the knee. People had to have a summed pain score of 125 to 400 on the index (more symptomatic) knee according to the WOMAC index and to be American Rheumatism Association functional class I-III
Exclusion criteria	Concurrent medical or arthritic conditions that could confound evaluation of the index joint; predominant patellofemoral disease; a history of clinically significant trauma or surgery to the index knee; coexisting disease that could preclude successful completion of the trial
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 58.6 (10.4). Gender (M:F): 568:1015. Ethnicity: White = 1239, Black = 221, Other = 123
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3. American Rheumatism Association functional class I-III. Duration of symptoms (mean [SD]): 10.0 (9.8) years.
Indirectness of population	No indirectness
Interventions	(n=317) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine hydrochloride 500mg three times daily (with

	double dummy placebo for other interventions). Duration 24 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg per day, except during the 24 hours before a clinical evaluation for joint pain. Other analgesics, including narcotics and NSAIDs were not permitted.. Indirectness: No indirectness (n=635) Intervention 2: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine plus chondroitin sulfate (500mg/400mg) taken three times a day or 400mg chondroitin sulfate taken three times a day (with double dummy placebo for other interventions). Duration 24 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg per day, except during the 24 hours before a clinical evaluation for joint pain. Other analgesics, including narcotics and NSAIDs were not permitted.. Indirectness: No indirectness Comments: These groups were not included interventions in the protocol and so were not included in the analysis. They are reported here for completeness. (n=318) Intervention 3: Specific COX-2 inhibitors (oral) - Celecoxib. Celecoxib 200mg daily (with double dummy placebo for other interventions). Duration 24 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg per day, except during the 24 hours before a clinical evaluation for joint pain. Other analgesics, including narcotics and NSAIDs were not permitted.. Indirectness: No indirectness (n=313) Intervention 4: Placebo. Placebo (double dummy to correspond to the other medications). Duration 24 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg per day, except during the 24 hours before a clinical evaluation for joint pain. Other analgesics, including narcotics and NSAIDs were not permitted.. Indirectness: No indirectness
Funding	Study funded by industry (Sponsored by the US Department of Veterans Affairs)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus CELECOXIB

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain score at 24 weeks; Group 1: mean -82.9 (SD 115.4); n=317, Group 2: mean -100 (SD 102.9); n=318; WOMAC pain score 0-500 Top=High is poor outcome; Comments: Reported glucosamine: -82.9 (115.4). Reported celecoxib: -100.0 (102.9). Baseline glucosamine: 233.3 (74.8). Baseline celecoxib: 234.9 (74.3).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis

symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function score at 24 weeks; Group 1: mean -222.3 (SD 388.3); n=317, Group 2: mean -289.3 (SD 340.7); n=318; WOMAC function score 0-1700 Top=High is poor outcome; Comments: Reported glucosamine: -222.3 (388.3). Reported celecoxib: -289.3 (340.7). Baseline glucosamine: 760.8 (328.2). Baseline celecoxib: 788.2 (309.0).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Congestive heart failure, stroke and chest pain at 24 weeks; Group 1: 1/317, Group 2: 1/318; Comments: Glucosamine: chest pain = 1. Celecoxib: stroke = 1.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only includes serious adverse events as there was incomplete reporting of all adverse events; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain score at 24 weeks; Group 1: mean -82.9 (SD 115.4); n=317, Group 2: mean -86.1 (SD 114.2); n=313; WOMAC pain score 0-500 Top=High is poor outcome; Comments: Reported glucosamine: -82.9 (115.4). Reported placebo: -86.1 (114.2). Baseline glucosamine: 233.3 (74.8). Baseline placebo: 237.1 (74.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function score at 24 weeks; Group 1: mean -222.3 (SD 388.3); n=317, Group 2: mean -227.4 (SD 362.7); n=313;

WOMAC function score 0-1700 Top=High is poor outcome; Comments: Reported glucosamine: -222.3 (388.3). Reported placebo: -227.4 (362.7). Baseline glucosamine: 760.8 (328.2). Baseline placebo: 765.8 (312.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Congestive heart failure, stroke and chest pain at 24 weeks; Group 1: 1/317, Group 2: 0/313; Comments: Glucosamine: Chest pain = 1
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only includes serious adverse events as there was incomplete reporting of all adverse events; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 75, Reason: Glucosamine: 75 withdrew. 9 adverse events, 27 lack of efficacy, 20 lost to follow-up, 10 other reasons.; Group 2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain score at 24 weeks; Group 1: mean -100 (SD 102.9); n=318, Group 2: mean -86.1 (SD 114.2); n=313; WOMAC pain score 0-500 Top=High is poor outcome; Comments: Reported celecoxib: -100.0 (102.9). Reported placebo: -86.1 (114.2). Baseline celecoxib: 234.9 (74.3). Baseline placebo: 237.1 (74.2).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.; Group 2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function score at 24 weeks; Group 1: mean -289.3 (SD 340.7); n=318, Group 2: mean -227.4 (SD 362.7); n=313; WOMAC function score 0-1700 Top=High is poor outcome; Comments: Reported celecoxib: -289.3 (340.7). Reported placebo: -227.4 (362.7). Baseline celecoxib: 788.2 (309.0). Baseline placebo: 765.8 (312.2).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.; Group

2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons. Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Congestive heart failure, stroke and chest pain at 24 weeks; Group 1: 1/318, Group 2: 0/313; Comments: Celecoxib: stroke = 1. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only includes serious adverse events as there was incomplete reporting of all adverse events; Baseline details: Reports age, sex, race, BMI, duration of osteoarthritis symptoms, time since diagnosis, number of paracetamol tablets per day, ARA functional class, radiographic severity, SF-36 score and baseline values of outcomes; Group 1 Number missing: 52, Reason: Celecoxib: 52 withdrew. 7 adverse events, 11 lack of efficacy, 17 lost to follow up, 17 other reasons.; Group 2 Number missing: 65, Reason: Placebo: 65 withdrew. 11 adverse events, 22 lack of efficacy, 17 lost to follow up, 15 other reasons.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Gana 2006[77] (Florete 2008[71], Kosinski 2007[107])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1011)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographically confirmed American College of Rheumatology functional class I-III osteoarthritis of the knee or hip
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women 18-74 years of age with radiographically confirmed American College of Rheumatology Functional Class I-III osteoarthritis of the knee or hip. They were required to have taken paracetamol, an NSAID, a COX-2 inhibitor or an opioid for at least 75 of the previous 90 days to treat osteoarthritis pain in the most painful hip or knee.
Exclusion criteria	Any medical condition other than osteoarthritis that was not well controlled; any other form of arthritis or joint disease at the index joint; a chronic pain syndrome or fibromyalgia; any contraindication for the use of tramadol; a history of substance abuse in the previous 6 months; any any condition that was likely to influence the absorption, efficacy, or safety of tramadol ER; people were not permitted to take another investigational medication, a corticosteroid, a medication that could interact with tramadol (e.g. carbamazepine), or another medication for pain (e.g., analgesics, antidepressants) during the study.
Recruitment/selection of patients	After a washout of previous analgesic treatment for 2-7 days, subjects were required to have a baseline index joint pain of ≥40mm on a 100-mm pain visual analogue scale (0 = no pain, 100 = extreme pain)
Age, gender and ethnicity	Age - Mean (SD): 58.2 (10.0). Gender (M:F): 380:631. Ethnicity: White = 791, Black = 174, Hispanic = 32, Asian = 8, Other = 6
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip (majority knee)).
Extra comments	Severity: Functional class I-III, mean class II Duration of symptoms (mean [SD]): 7.8 (7.3) years

Indirectness of population	No indirectness
Interventions	(n=806) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol ER 100mg, 200mg, 300mg or 400mg once daily (with either no to three placebo tablets to blind the participants about the dose of tramadol they were receiving so that all participants had four tablets once per day). Duration 12 weeks. Concurrent medication/care: People could take up to 200mg/day of paracetamol for no more than 3 consecutive days for reasons other than osteoarthritis of chronic pain. The use of paracetamol was prohibited during the washout period and in the 48 hours before each study visit after the screening visit.. Indirectness: No indirectness Comments: Each dose of tramadol was reported separately but combined for analysis due to class effect as agreed in the protocol (n=205) Intervention 2: Placebo. Four placebo tablets once per day. Duration 12 weeks. Concurrent medication/care: People could take up to 200mg/day of paracetamol for no more than 3 consecutive days for reasons other than osteoarthritis of chronic pain. The use of paracetamol was prohibited during the washout period and in the 48 hours before each study visit after the screening visit.. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a grant from Biovail Laboratories International SRL.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical component at 12 weeks; Group 1: mean 3.6 (SD 8.5); n=806, Group 2: mean 2.4 (SD 8.6); n=205; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Calculated SD from this. Reported tramadol 100mg: 3.6 (0.6). Reported tramadol 200mg: 3.9 (0.6). Reported tramadol 300mg: 3.6 (0.6). Reported tramadol 400mg: 3.2 (0.6). Reported placebo: 2.4 (0.6). Baseline values not reported.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, height, weight, index joint, osteoarthritis duration, functional class, baseline values for WOMAC pain and physical function, global assessment of disease activity, quality of sleep and pain intensity from a daily diary. Does not report baseline values for SF-36.; Group 1 Number missing: 363, Reason: Tramadol 100mg: 203 randomised. 1 did not receive the treatment. 29 had adverse events. 31 had lack of efficacy. 11 subject choice. 11 other. Tramadol 200mg: 203 randomised. 2 did not receive treatment. 40 adverse events. 29 lack of efficacy. 6 subject choice. 10 other. Tramadol 300mg: 204 randomised. 3 did not receive treatment. 54 adverse events. 18 lack of efficacy. 14 subject choice. 11 other. Tramadol 400mg: 205 randomised. 3 did not receive treatment. 60 adverse events. 23 lack of efficacy. 8 subject choice. 8 other.; Group 2 Number missing: 90, Reason: 205 randomised. 21 adverse events. 46 lack of efficacy. 9 subject choice. 11 other.

- Actual outcome for Other: SF-36 mental component at 12 weeks; Group 1: mean 0.1 (SD 8.5); n=806, Group 2: mean -0.3 (SD 8.6); n=205; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Calculated SD from this. Reported tramadol 100mg: 1.1 (0.6). Reported tramadol 200mg: 0.6 (0.6). Reported tramadol 300mg: -0.7 (0.6). Reported tramadol 400mg: -0.5 (0.6). Reported placebo: -0.3 (0.6). Baseline values not reported.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, height, weight, index joint, osteoarthritis duration, functional class, baseline values for WOMAC pain and physical function, global assessment of disease activity, quality of sleep and pain intensity from a daily diary. Does not report baseline values for SF-36.; Group 1 Number missing: 363, Reason: Tramadol 100mg: 203 randomised. 1 did not receive the treatment. 29 had adverse events. 31 had lack of efficacy. 11 subject choice. 11 other. Tramadol 200mg: 203 randomised. 2 did not receive treatment. 40 adverse events. 29 lack of efficacy. 6 subject choice. 10 other. Tramadol 300mg: 204 randomised. 3 did not receive treatment. 54 adverse events. 18 lack of efficacy. 14 subject choice. 11 other. Tramadol 400mg: 205 randomised. 3 did not receive treatment. 60 adverse events. 23 lack of efficacy. 8 subject choice. 8 other.; Group 2 Number missing: 90, Reason: 205 randomised. 21 adverse events. 46 lack of efficacy. 9 subject choice. 11 other.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean 107.6 (SD 123.2); n=806, Group 2: mean 74.2 (SD 123.1); n=205; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Reports least square means and standard error. Calculated SD from this. Reported tramadol 100mg: 107.2 (8.6). Reported tramadol 200mg: 111.5 (8.7). Reported tramadol 300mg: 103.9 (8.7). Reported tramadol 400mg: 107.8 (8.7). Reported placebo: 74.2 (8.6). Baseline (mean [SD]) tramadol 100mg: 308.2 (99.3). Baseline tramadol 200mg: 315.2 (92.4). Baseline tramadol 300mg: 296.6 (96.3). Baseline tramadol 400mg: 298.0 (93.7). Baseline placebo: 305.9 (95.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, height, weight, index joint, osteoarthritis duration, functional class, baseline values for WOMAC pain and physical function, global assessment of disease activity, quality of sleep and pain intensity from a daily diary. Does not report baseline values for SF-36.; Group 1 Number missing: 363, Reason: Tramadol 100mg: 203 randomised. 1 did not receive the treatment. 29 had adverse events. 31 had lack of efficacy. 11 subject choice. 11 other. Tramadol 200mg: 203 randomised. 2 did not receive treatment. 40 adverse events. 29 lack of efficacy. 6 subject choice. 10 other. Tramadol 300mg: 204 randomised. 3 did not receive treatment. 54 adverse events. 18 lack of efficacy. 14 subject choice. 11 other. Tramadol 400mg: 205 randomised. 3 did not receive treatment. 60 adverse events. 23 lack of efficacy. 8 subject choice. 8 other.; Group 2 Number missing: 90, Reason: 205 randomised. 21 adverse events. 46 lack of efficacy. 9 subject choice. 11 other.

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean 336.9 (SD 408.6); n=806, Group 2: mean 234.3 (SD 402.3); n=205; WOMAC physical function 0-1700 Top=High is poor outcome; Comments: Reports least square means and standard error. Calculated SD from this. Reported tramadol 100mg: 331.7 (28.5). Reported tramadol 200mg: 350.2 (29.0). Reported tramadol 300mg: 336.1 (28.8). Reported tramadol 400mg: 329.8 (28.8). Reported placebo: 234.3 (28.1). Baseline (mean [SD]) tramadol 100mg: 1071.6 (331.2). Baseline tramadol 200mg: 1096.2 (298.7). Baseline tramadol 300mg: 1026.6 (337.6). Baseline tramadol 400mg: 1010.9 (331.7). Baseline placebo: 1058.7 (340.3).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, height, weight, index joint, osteoarthritis duration, functional class, baseline values for WOMAC pain and physical function, global assessment of disease activity, quality of sleep and pain intensity from a daily diary. Does not report baseline values for SF-36.; Group 1 Number missing: 363, Reason: Tramadol 100mg: 203 randomised. 1 did not

receive the treatment. 29 had adverse events. 31 had lack of efficacy. 11 subject choice. 11 other. Tramadol 200mg: 203 randomised. 2 did not receive treatment. 40 adverse events. 29 lack of efficacy. 6 subject choice. 10 other. Tramadol 300mg: 204 randomised. 3 did not receive treatment. 54 adverse events. 18 lack of efficacy. 14 subject choice. 11 other. Tramadol 400mg: 205 randomised. 3 did not receive treatment. 60 adverse events. 23 lack of efficacy. 8 subject choice. 8 other.; Group 2 Number missing: 90, Reason: 205 randomised. 21 adverse events. 46 lack of efficacy. 9 subject choice. 11 other. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Postural hypotension at 12 weeks; Group 1: 24/806, Group 2: 6/205; Comments: Tramadol 100mg = 1, tramadol 200mg = 9, tramadol 300mg = 3, tramadol 400mg = 11, placebo = 6 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, height, weight, index joint, osteoarthritis duration, functional class, baseline values for WOMAC pain and physical function, global assessment of disease activity, quality of sleep and pain intensity from a daily diary. Does not report baseline values for SF-36.; Group 1 Number missing: 363, Reason: Tramadol 100mg: 203 randomised. 1 did not receive the treatment. 29 had adverse events. 31 had lack of efficacy. 11 subject choice. 11 other. Tramadol 200mg: 203 randomised. 2 did not receive treatment. 40 adverse events. 29 lack of efficacy. 6 subject choice. 10 other. Tramadol 300mg: 204 randomised. 3 did not receive treatment. 54 adverse events. 18 lack of efficacy. 14 subject choice. 11 other. Tramadol 400mg: 205 randomised. 3 did not receive treatment. 60 adverse events. 23 lack of efficacy. 8 subject choice. 8 other.; Group 2 Number missing: 90, Reason: 205 randomised. 21 adverse events. 46 lack of efficacy. 9 subject choice. 11 other.	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Geba 2002[78]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=382)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee for at least 6 months who fulfilled the American College of Rheumatology clinical criteria for osteoarthritis of the knee and had an American College of Rheumatology functional class rating of I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and non-pregnant women with symptomatic osteoarthritis of the knee for at least 6 months if they were at least 40 years old; fulfilled the American College of Rheumatology clinical criteria for osteoarthritis of the knee; and had an American College of Rheumatology functional class rating of I-III. Previous users of either a single, prescription-strength NSAID or high doses of paracetamol for control of osteoarthritis symptoms for at least 30 days prior to entry. entry criteria for NSAID users included demonstrating, while taking the NSAID, a screening WOMAC visual analogue scale score of less than 80mm for the assessment of walking pain at visit 1 (screening); a minimum VAS score of 40mm for the assessment of walking pain after discontinuing the NSAID; an increase (ie. worsening) from screening in the walking pain VAS score of at least 15mm, and a worsening from screening in the Investigator Global Assessment of Disease Status of at least 1 point on a 5-point Likert Scale at visit 2. During the prespecified washout period, people who had discontinued NSAIDs could take paracetamol for intolerable pain, but paracetamol was discontinued for at least 24 hours before efficacy assessments. Since previous paracetamol users were not taking NSAIDs they were required to have a history of therapeutic benefit with regular doses of paracetamol (1200-4000mg/day) as exclusive therapy of osteoarthritis, a score of 2-4 on IGADS and the same minimum VAS score of 40mm required of the NSAID previous users after discontinuing paracetamol therapy for the assessment of walking pain at visits 1 and 2.
Exclusion criteria	Concurrent medical or arthritis disease or abnormal laboratory results that had potential to confound or interfere with the efficacy evaluation or pose an additional risk to the person; history of allergy to study drugs, hypersensitivity to aspirin, ibuprofen, or

	any other NSAID or sulfnamide containing compounds; or received an investigational drug within 30 days of screening
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.6 (10.81). Gender (M:F): 121:261. Ethnicity: White = 326, Hispanic = 27, Black = 25, Other = 4
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: American College of Rheumatology functional class rating of I-III Duration of symptoms: Not stated explicitly. At least 6 months..
Indirectness of population	No indirectness
Interventions	(n=97) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg per day and matching placebo tablets. Duration 6 weeks. Concurrent medication/care: NSAID users were allowed to take 325mg paracetamol tablets during the washout phase as rescue therapy for osteoarthritis pain at a daily dose of paracetamol restricted to 2600mg per day. No other rescue medication was allowed during the study. All people discontinued use of paracetamol at least 12 hours before visit 1 and visit 2. Indirectness: No indirectness (n=191) Intervention 2: NSAIDs - Other. Rofecoxib 12.5mg or 25mg once a day and matching placebo tablets. Duration 6 weeks. Concurrent medication/care: NSAID users were allowed to take 325mg paracetamol tablets during the washout phase as rescue therapy for osteoarthritis pain at a daily dose of paracetamol restricted to 2600mg per day. No other rescue medication was allowed during the study. All people discontinued use of paracetamol at least 12 hours before visit 1 and visit 2. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=94) Intervention 3: Paracetemol (oral) - Paracetemol. Paracetamol 4000mg per day - 1000mg four times a day and matching placebo tablets. Duration 6 weeks. Concurrent medication/care: NSAID users were allowed to take 325mg paracetamol tablets during the washout phase as rescue therapy for osteoarthritis pain at a daily dose of paracetamol restricted to 2600mg per day. No other rescue medication was allowed during the study. All people discontinued use of paracetamol at least 12 hours before visit 1 and visit 2. Indirectness: No indirectness

Funding	Study funded by industry (Research was supported by Merck & Co, Inc.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PARACETEMOL Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -28.6 (SD 22.8); n=94, Group 2: mean -24.9 (SD 22.5); n=92; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported celecoxib: -28.6 (-33.2 to -24.0). Reported paracetamol: -24.9 (-29.5 to -20.3). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity and prior drug use. Does not report baseline WOMAC subscales values; Group 1 Number missing: 17, Reason: 9 lack of efficacy, 4 adverse events, 4 others; Group 2 Number missing: 29, Reason: 16 lack of efficacy, 1 lost to follow up, 6 adverse events, 6 others Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC function subscale at 6 weeks; Group 1: mean -24.9 (SD 21.8); n=94, Group 2: mean -19.5 (SD 22.5); n=92; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported celecoxib: -24.9 (-29.3 to -20.5). Reported paracetamol: -19.5 (-24.1 to -14.9). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity and prior drug use. Does not report baseline WOMAC subscales values; Group 1 Number missing: 17, Reason: 9 lack of efficacy, 4 adverse events, 4 others; Group 2 Number missing: 29, Reason: 16 lack of efficacy, 1 lost to follow up, 6 adverse events, 6 others	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Ghosh 2007[79]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=585)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis having symptoms for at least 6 months with radiological evidence
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 40-64 years, of either sex, with knee osteoarthritis and those who met American Rheumatologic Association functional class I-III.
Exclusion criteria	Significant renal impairment (creatinine clearance <30mL/minute); major cardiac or hepatic diseases; bronchial asthma; any active gastro-intestinal bleeding; neoplasia or acute meniscus injury; arthroscopy in the study joint in the last 6 months; obesity (>100kg); allergic to conventional NSAIDs; people who required systematic steroids, warfarin, lithium, low dose aspirin, anti-ulcer drugs or intra-articular steroids within the last 2 months.
Recruitment/selection of patients	People who entered the study underwent a pretreatment washout period for 10 days when baseline investigations were done. Does not state that people were excluded for any changes in baseline investigations.
Age, gender and ethnicity	Age - Mean (range): 54.9 (40-64). Gender (M:F): 217:368. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: American Rheumatologic Association Functional Class I-III Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=304) Intervention 1: NSAIDs - Diclofenac. Diclofenac sustained release preparation (75mg once daily at bed time) or Etoricoxib 90mg once daily at bed time. Duration 4 weeks. Concurrent medication/care: Paracetamol (up to 2g/day for a maximum of 3 days but not before assessment of arthritis) was permitted. Indirectness: No indirectness Comments: The diclofenac and etoricoxib groups were combined for class effect as

	agreed in the protocol (n=123) Intervention 2: Placebo. Matching placebo (once daily at bed time). Duration 4 weeks. Concurrent medication/care: Paracetamol (up to 2g/day for a maximum of 3 days but not before assessment of arthritis) was permitted. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC/ETORICOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain (visual analogue scale) at 4 weeks; Group 1: mean -4.3 (SD 2.3); n=304, Group 2: mean -0.82 (SD 1.3); n=123; VAS 0-10 Top=High is poor outcome; Comments: Reports mean change scores and standard error. Calculated SD from this. Reported etoricoxib: 4.69 (0.18). Reported diclofenac: 3.96 (0.19). Reported placebo: 0.82 (0.12). Calculated SD etoricoxib: 2.3. Calculated SD diclofenac: 2.3. Calculated SD placebo: 1.3.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Reports that "the code was broken once only at the end of the 2nd week to eliminate any 'real' fatal outcome"; Indirectness of outcome: No indirectness ; Baseline details: Reports overall age and gender. Does not report any baseline values for outcomes.; Group 1 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm; Group 2 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal side effects at 4 weeks; Group 1: 33/304, Group 2: 4/123
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Reports that "the code was broken once only at the end of the 2nd week to eliminate any 'real' fatal outcome"; Indirectness of outcome: No indirectness ; Baseline details: Reports overall age and gender. Does not report any baseline values for outcomes.; Group 1 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm; Group 2 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm
- Actual outcome for Knee: Haematemesis at 4 weeks; Group 1: 1/304, Group 2: 0/123
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Reports that "the code was broken once only at the end of the 2nd week to eliminate any 'real' fatal outcome"; Indirectness of outcome: No indirectness ; Baseline details: Reports overall age and gender. Does not report any baseline values for outcomes.; Group 1 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm; Group 2 Number missing: 0, Reason:

Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Cerebrovascular accident at 4 weeks; Group 1: 1/304, Group 2: 0/123 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Reports that "the code was broken once only at the end of the 2nd week to eliminate any 'real' fatal outcome"; Indirectness of outcome: No indirectness ; Baseline details: Reports overall age and gender. Does not report any baseline values for outcomes.; Group 1 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm; Group 2 Number missing: 0, Reason: Only reports overall numbers. Reports that 142 (72.82%) of people completed the study in the diclofenac arm, 162 (93.08%) completed the study in the etoricoxib arm and 123 (63.08%) completed the study in the placebo arm	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Giansiracusa 1977[80]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=437)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (4 weeks per intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a clinical diagnosis of degenerative joint disease of six months or more duration which involved one or more major joints. Roentgenograms confirmed the diagnosis in most cases.
Stratum	Multisite:
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a clinical diagnosis of degenerative joint disease of six months or more duration which involved one or more major joints.
Exclusion criteria	People with serious complicating disease or other forms of arthritis, people receiving treatment such as phenylbutazone or corticosteroids which might complicate or confound interpretation of results
Recruitment/selection of patients	Multiclinic study
Age, gender and ethnicity	Age - Other: Mean: 64.4. Gender (M:F): 124:313. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee, hip and shoulder).
Extra comments	Severity: Not stated Duration of symptoms (mean): 11.6 years.
Indirectness of population	No indirectness
Interventions	(n=437) Intervention 1: NSAIDs - Ibuprofen. Either ibuprofen 150mg four times daily or aspirin 300mg four times daily. Duration 4 weeks. Concurrent medication/care: People were instructed not to use aspirin or other analgesic antiarthritic medications. They supplied propoxyphene, which has no known anti-inflammatory activity, as an adjunctive analgesic for people whose pain was not adequately controlled by the study medication. Indirectness: No indirectness (n=437) Intervention 2: Placebo. Matching placebo four times daily. Duration 4 weeks. Concurrent medication/care: People were instructed not to use aspirin or other

	analgesic antiarthritic medications. They supplied propoxyphene, which has no known anti-inflammatory activity, as an adjunctive analgesic for people whose pain was not adequately controlled by the study medication. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by grants from The Upjohn Company)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN AND ASPIRIN versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastric ulcer at 4 weeks; Group 1: 1/437, Group 2: 0/437; Comments: Ibuprofen = 0, aspirin = 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, mean duration of arthritis, mean body weight, and joint involved; Group 1 Number missing: 51, Reason: 378 completed the trial. 7 withdrew due to aspirin. 5 due to placebo. 1 due to ibuprofen.; Group 2 Number missing: 54, Reason: 378 completed the trial. 7 withdrew due to aspirin. 5 due to placebo. 1 due to ibuprofen.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Gibofsky 2003[82]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=477)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee, diagnosed according to the American College of Rheumatology criteria with: a functional capacity class rating of I-III; osteoarthritis in a flare state at baseline; and a negative serum of urine pregnancy test (for women of childbearing potential).
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female adults, ≥40 years old, with osteoarthritis of the knee, diagnosed according to the American College of Rheumatology criteria with: a functional capacity class rating of I-III; osteoarthritis in a flare state at baseline; and a negative serum of urine pregnancy test (for women of childbearing potential). People who had previously received NSAIDs or COX-2 specific inhibitors were eligible for enrollment.
Exclusion criteria	Any inflammatory arthritis or acute joint trauma in the index knee; recent corticosteroid (previous 8 weeks) or hyaluronic acid (previous 6 months) infection; NSAID use (with the exception of aspirin ≤325mg/day, for an indication other than arthritis) within the previous 2 days or 5 half-lives of the drug; active malignancy or a history of malignancy (unless the lesion was surgically removed, with no evidence of recurrence for at least 5 years); upper gastrointestinal ulceration within the past 30 days; active GI disease; chronic or acute renal or hepatic disorder; significant coagulation defect; known hypersensitivity to NSAIDs or COX-2 specific inhibitors; or abnormal laboratory test results at the screening visit
Recruitment/selection of patients	For people with uncontrolled osteoarthritis who were not receiving an NSAIDs or analgesic, 3 of the following 4 flare criteria were required to be met based on the pre-treatment assessments: VAS score of more than or equal to 40mm for patient's assessment of OA pain; OA severity index of more than or equal to 7; patient's global assessment of arthritis of "poor" or "very poor"; a physician's global assessment of arthritis of "poor" or "very poor". People meeting these criteria could be enrolled at the screening visit without the need for a separate baseline visit. People taking NSAID or analgesic therapy were instructed to discontinue and return when their symptoms worsened (2-14 days after the screening visit). Arthritis flare was

	defined as patient's and physician's global assessments of "fair", "poor" or "very poor", and any 3 of the following 4 criteria: worsening of at least 1 grade from screening in the patient's global assessment, worsening of at least 1 grade from screening in the physician's global assessment, patient's assessment of OA pain on the VAS of more than or equal to 40mm, or worsening in the OA severity index of more than or equal to 2 points.
Age, gender and ethnicity	Age - Mean (SD): 62.9 (10.3). Gender (M:F): 157:320. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity class rating of I-III Duration of symptoms (mean [SD]): 8.6 (8.1) years.
Indirectness of population	No indirectness
Interventions	(n=189) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day with a placebo tablet once per day. Duration 6 weeks. Concurrent medication/care: During the study period, people were not allowed to take NSAIDs, analgesics, oral or injectable corticosteroids or hyaluronic acid, anticoagulants, disease-modifying antirheumatic drugs, anti-ulcer medication, daily or almost daily use of antacids, or antiplatelet agents. However, use of low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis was allowed, as was occasional use of paracetamol (use of paracetamol had to be discontinued for 48 hours prior to the arthritis assessments) or antacids. Indirectness: No indirectness (n=190) Intervention 2: NSAIDs - Other. Rofecoxib 25mg once a day with a placebo tablet once a day. Duration 6 weeks. Concurrent medication/care: During the study period, people were not allowed to take NSAIDs, analgesics, oral or injectable corticosteroids or hyaluronic acid, anticoagulants, disease-modifying antirheumatic drugs, anti-ulcer medication, daily or almost daily use of antacids, or antiplatelet agents. However, use of low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis was allowed, as was occasional use of paracetamol (use of paracetamol had to be discontinued for 48 hours prior to the arthritis assessments) or antacids. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=96) Intervention 3: Placebo. Two placebo tablets once a day. Duration 6 weeks. Concurrent medication/care: During the study period, people were not allowed to take NSAIDs, analgesics, oral or injectable corticosteroids or hyaluronic acid,

	anticoagulants, disease-modifying antirheumatic drugs, anti-ulcer medication, daily or almost daily use of antacids, or antiplatelet agents. However, use of low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis was allowed, as was occasional use of paracetamol (use of paracetamol had to be discontinued for 48 hours prior to the arthritis assessments) or antacids. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Pharmacia Corporation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -4.7 (SD 4.1); n=189, Group 2: mean -2.6 (SD 3.9); n=96; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported celecoxib: -4.7 (0.3). Reported placebo: -2.6 (0.4). Baseline celecoxib: 11.2 (0.3). Baseline placebo: 11.0 (0.3).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, OA pain score, WOMAC scores, patient's and physician's global assessment and history of gastrointestinal adverse events; Group 1 Number missing: 31, Reason: 31 withdrew. 1 lost to follow up, 3 preexisting violations, 4 protocol noncompliance, 10 treatment failure, 11 adverse events, 2 others.; Group 2 Number missing: 32, Reason: 34 withdrew (including 2 people not in the ITT cohort - total population number = 98). 1 lost to follow-up, 2 preexisting violations, 2 protocol noncompliance, 21 treatment failure, 5 adverse events, 3 others.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical functioning subscale at 6 weeks; Group 1: mean -14.7 (SD 13.7); n=189, Group 2: mean -8.2 (SD 12.7); n=96; WOMAC physical functioning subscale 0-68 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported celecoxib: -14.7 (1.0). Reported placebo: -8.2 (1.3). Baseline celecoxib: 38.8 (0.8). Baseline placebo: 38.4 (1.0).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, OA pain score, WOMAC scores, patient's and physician's global assessment and history of gastrointestinal adverse events; Group 1 Number missing: 31, Reason: 31 withdrew. 1 lost to follow up, 3 preexisting violations, 4 protocol noncompliance, 10 treatment failure, 11 adverse events, 2 others.; Group 2 Number missing: 32, Reason: 34 withdrew (including 2 people not in the ITT cohort - total population number = 98). 1 lost to follow-up, 2 preexisting violations, 2 protocol noncompliance, 21 treatment failure, 5 adverse events, 3 others.

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2:

	Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Gibofsky 2014[81]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=305)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinically and radiographically confirmed hip and/or knee osteoarthritis (Kellgren-Lawrence grade II-III).
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women with clinically and radiographically confirmed hip and/or knee osteoarthritis (Kellgren-Lawrence grade II-III) who were ≥40 years of age with a body weight ≥45kg and a body mass index ≤40kg/m². Enrolled people had WOMAC pain subscale scores of ≥40mm at baseline (day 0 prior to the first dose of the study drug) by an 100mm visual analogue scale and were required to have experienced a documented OA pain flare (≥15mm increase in WOMAC pain subscale score from screening to baseline) following NSAID or paracetamol discontinuation.
Exclusion criteria	Women of childbearing potential if pregnant or lactating or not taking a medically acceptable method of birth control; history of an allergy to NSAIDs or paracetamol; clinically significant unstable medical condition; a history of alcohol or drug abuse; any arthritic or painful disease or major surgery in the target joint (e.g. joint replacement) that could confound or interfere with the evaluation of efficacy; aspartate aminotransferase or alanine aminotransferase ≥3x the upper limit of normal at screening
Recruitment/selection of patients	People reporting chronic NSAID and/or paracetamol use to treat osteoarthritis pain were instructed to discontinue taking analgesic medications for a minimum of 5 days prior to study entry. They were required to have experienced a documented OA pain flare (≥15mm increase in WOMAC pain subscale score from screening to baseline) following NSAID or paracetamol discontinuation.
Age, gender and ethnicity	Age - Mean (SD): 61.6 (8.9). Gender (M:F): 102:203. Ethnicity: Native American or Alaskan Native = 1, Native Black or African American = 57, Native Hawaiian or other Pacific Islander = 1, White = 245, Other = 2

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Kellgren-Lawrence grade II-III Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=202) Intervention 1: NSAIDs - Diclofenac. Low-dose submicron diclofenac capsules 35mg twice or three times daily with placebo to make four dosing periods each day. Duration 12 weeks. Concurrent medication/care: Paracetamol (up to 500mg) every 4 to 6 hours as required was permitted as rescue analgesia (to not exceed 3000mg daily).. Indirectness: No indirectness Comments: The results for the two different dose regimens of diclofenac were combined for class effect as agreed in the protocol. (n=103) Intervention 2: Placebo. Placebo four times daily. Duration 12 weeks. Concurrent medication/care: Paracetamol (up to 500mg) every 4 to 6 hours as required was permitted as rescue analgesia (to not exceed 3000mg daily).. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by Iroko Pharmaceuticals LLC, Philadelphia, PA, USA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -41.5 (SD 30.2); n=202, Group 2: mean -32.5 (SD 29.8); n=103; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports mean change score and SE. SD calculated from this. Diclofenac groups were combined. Reported diclofenac 35mg TDS = -44.1 (3.07). Reported diclofenac 35mg BD: -39.0 (2.91). Reported placebo: -32.5 (2.94). Calculated SD diclofenac 35mg TDS = 30.4. Calculated SD diclofenac 35mg BD = 29.7. Calculated SD placebo = 29.8.
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, target osteoarthritis joint and radiographic severity, race, age, weight, height and BMI. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 30, Reason: Adverse events = 21, lack of efficacy = 2, withdrew consent = 4, protocol violation = 3; Group 2 Number missing: 17, Reason: Adverse event = 4, lack of efficacy = 6, withdrew consent = 2, protocol violation = 2, lost to follow-up = 3

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC function subscale at 12 weeks; Group 1: mean -33 (SD 27.4); n=202, Group 2: mean -23.2 (SD 27); n=103; WOMAC

function subscale 0-100 Top=High is poor outcome; Comments: Reports mean change score and SE. SD calculated from this. Diclofenac groups were combined. Reported diclofenac 35mg TDS = -35.9 (2.80). Reported diclofenac 35mg BD: -30.3 (2.63). Reported placebo: -23.2 (2.66). Calculated SD diclofenac 35mg TDS = 27.7. Calculated SD diclofenac 35mg BD = 26.8. Calculated SD placebo = 27.0. Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, target osteoarthritis joint and radiographic severity, race, age, weight, height and BMI. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 30, Reason: Adverse events = 21, lack of efficacy = 2, withdrew consent = 4, protocol violation = 3; Group 2 Number missing: 17, Reason: Adverse event = 4, lack of efficacy = 6, withdrew consent = 2, protocol violation = 2, lost to follow-up = 3	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Giordano 2009[83]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Italy; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Other: Intervention and follow up for 12 weeks. Additional follow up (without intervention) for 12 weeks.
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of primary monolateral or bilateral knee osteoarthritis who met the American Rheumatism Association criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with both monolateral or bilateral knee osteoarthritis who met the American Rheumatism Association criteria. Symptomatic for at least 3 months before enrollment and have a radiographic grade between I and III, as measured with the Kellgren-Lawrence method.
Exclusion criteria	Haematologic disorders; renal disease; liver disease; diabetes mellitus; acute illness; neoplasms; other rheumatic diseases; disabling comorbid conditions that would make it impossible for the person to visit the research center; pregnancy or nursing; body mass index >30kg/m²; people with grade IV osteoarthritis; people who had had joint lavage, arthroscopy, or treatment with hyaluronic acid or other disease-modifying agents during the previous 6 months; or who had been treated with intra-articular corticosteroids during the past 3 months.
Recruitment/selection of patients	Consecutive outpatients.
Age, gender and ethnicity	Age - Mean (SD): 57.7 (7.8). Gender (M:F): 18:42. Ethnicity: All people were white
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade I-III (median II-III). Duration of disease (mean [SD]): 6.3 (4.8) years.
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 1500mg once a day for 12 weeks as sachets of power for oral solution. Duration 12 weeks. Concurrent medication/care: During the study, it was recommended that people do not modify their therapeutic

	program (for both drug treatments and physical therapy) unless an adverse event occurred and required management. In particular, they were instructed to avoid corticosteroids and hyaluronic acid infiltrations, arthroscopic surgery and joint lavage, and to avoid treatment with disease modifying osteoarthritis drugs. For rescue analgesia, people were allowed paracetamol 500mg, diclofenac 150mg, piroxicam 20mg, naproxen 750mg, or aceclofenac 200mg, all of which were to be used as needed and noted daily in a diary. Indirectness: No indirectness (n=30) Intervention 2: Placebo. Matching placebo sachet of powder. Duration 12 weeks. Concurrent medication/care: During the study, it was recommended that people do not modify their therapeutic program (for both drug treatments and physical therapy) unless an adverse event occurred and required management. In particular, they were instructed to avoid corticosteroids and hyaluronic acid infiltrations, arthroscopic surgery and joint lavage, and to avoid treatment with disease modifying osteoarthritis drugs. For rescue analgesia, people were allowed paracetamol 500mg, diclofenac 150mg, piroxicam 20mg, naproxen 750mg, or aceclofenac 200mg, all of which were to be used as needed and noted daily in a diary. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC total pain score at 12 weeks; Group 1: mean 30.56 (SD 11.5); n=30, Group 2: mean 53.3 (SD 7.1); n=30; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 51.2 (8.3). Baseline placebo: 50.03 (6.4).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, BMI, disease duration, Kellgren-Lawrence score, previous NSAID/paracetamol intake and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up; Group 2 Number missing: 2, Reason: 2 lost to follow up
- Actual outcome for Knee: WOMAC total pain score at 24 weeks; Group 1: mean 47.75 (SD 14.5); n=30, Group 2: mean 51.05 (SD 6.7); n=30; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 51.2 (8.3). Baseline placebo: 50.03 (6.4).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, BMI, disease duration, Kellgren-Lawrence score, previous NSAID/paracetamol intake and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up; Group 2 Number missing: 2, Reason: 2 lost to follow up

Protocol outcome 2: Physical function at ≤3- or >3- months

- Actual outcome for Knee: WOMAC total physical function score at 12 weeks; Group 1: mean 32.82 (SD 13.2); n=30, Group 2: mean 55.1 (SD 14.9); n=30; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 52.16 (12.3). Baseline placebo: 53.94 (14.1). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, BMI, disease duration, Kellgren-Lawrence score, previous NSAID/paracetamol intake and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up; Group 2 Number missing: 2, Reason: 2 lost to follow up - Actual outcome for Knee: WOMAC total physical function score at 24 weeks; Group 1: mean 51.85 (SD 12.5); n=30, Group 2: mean 53.27 (SD 14); n=30; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 52.16 (12.3). Baseline placebo: 53.94 (14.1). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, BMI, disease duration, Kellgren-Lawrence score, previous NSAID/paracetamol intake and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up; Group 2 Number missing: 2, Reason: 2 lost to follow up Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 12 weeks; Group 1: 1/30, Group 2: 2/30 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, BMI, disease duration, Kellgren-Lawrence score, previous NSAID/paracetamol intake and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up; Group 2 Number missing: 2, Reason: 2 lost to follow up	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Golden 2004[84]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (pooled analysis of 2 identical trials) (n=465)
Countries and setting	Conducted in Switzerland; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 7 days on the intervention, 8 days additional follow up after this
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: All people had radiographic changes including: subchondral sclerosis, joint space narrowing, presence of osteophytes or marginal lipping, or cyst formation in the knee joint
Stratum	Knee:
Subgroup analysis within study	Not applicable
Inclusion criteria	People in general good health, male or female, of any race and aged at least 25 years. All people were diagnosed with osteoarthritis that showed at least one of the following radiographic changes: subchondral sclerosis, joint space narrowing, presence of osteophytes or marginal lipping, or cyst formation in the knee joint of osteoarthritis stage I-III, as documented in the previous 3 years; all people had at least moderate pain in the knee on weight bearing as measured on a categorical scale, had episodic flares of osteoarthritis, and had completed the required washout period
Exclusion criteria	Stage IV osteoarthritis; moderate to severe chronic low back pain; inflammatory joint diseases including rheumatoid arthritis, gout, mixed connective tissue disease, seronegative spondyloarthropathy; psoriatic arthritis; systemic lupus erythematosus; daily regimen of prescription NSAIDs for arthritis pain for the past 3 months (occasional periodic use of NSAIDs was not prohibited); had a recent traumatic injury; had a history of hypersensitivity or intolerance to any of the study medications; history of peptic ulceration within the previous 9 months; gastrointestinal surgery; GI complaints or GI dysfunction that could interfere with drug absorption; any other significant medical conditions
Recruitment/selection of patients	2 different clinical trials with identical criteria/methodology (sponsored by the same company). Multicenter trial (10 centers).
Age, gender and ethnicity	Age - Mean (SD): 60.57 (12.78). Gender (M:F): 142:323. Ethnicity: 400 were Caucasian, 40 were Black, 2 were Asian, 22 were Hispanic, 1 was Other
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity of osteoarthritis: Functional class I-III. Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=162) Intervention 1: NSAIDs - Naproxen. 220mg naproxen sodium (oral) three times a day (if the person was age 65 years or over they were to take it two times a day). Duration 7 days. Concurrent medication/care: Paracetamol (500mg up to three times a day) could be taken during the 2 day/4 day washout period (if previously on short-term/long-term NSAIDs respectively) but could not be used during the study. Indirectness: No indirectness (n=148) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg four times daily (oral). Duration 7 days. Concurrent medication/care: Paracetamol (500mg up to three times a day) could be taken during the 2 day/4 day washout period (if previously on short-term/long-term NSAIDs respectively) but could not be used during the study. Indirectness: No indirectness (n=155) Intervention 3: Placebo. Placebo tablets four times a day. Duration 7 days. Concurrent medication/care: Paracetamol (500mg up to three times a day) could be taken during the 2 day/4 day washout period (if previously on short-term/long-term NSAIDs respectively) but could not be used during the study. Indirectness: No indirectness
Funding	Study funded by industry (These studies were sponsored by F. Hoffmann-La Roche AG, Basel, Switzerland)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PARACETEMOL

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders at 15 days; Group 1: 36/161, Group 2: 31/148; Comments: Provides what the adverse events were for some participants (but not all). Naproxen: Dyspepsia = 10, nausea = 8, diarrhoea = 4, constipation = 5, abdominal pain = 1, gastrointestinal upset = 3, vomiting = 2. Paracetamol: Dyspepsia = 11, nausea = 8, diarrhoea = 3, constipation = 1, abdominal pain = 1, gastrointestinal upset = 1, loose stools = 1, vomiting = 2, pharyngolarygeal pain = 2.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0
- Actual outcome for Knee: Rectal bleeding at 15 days; Group 1: 3/162, Group 2: 0/148
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cardiac disorders at 15 days; Group 1: 0/161, Group 2: 1/148; Comments: No information on what the events were
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nervous system disorders at 15 days; Group 1: 13/161, Group 2: 5/148; Comments: No information on what the events were
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders at 15 days; Group 1: 36/161, Group 2: 27/155; Comments: Provides what the adverse events were for some participants (but not all). Naproxen: Dyspepsia = 10, nausea = 8, diarrhoea = 4, constipation = 5, abdominal pain = 1, gastrointestinal upset = 3, vomiting = 2. Placebo: Dyspepsia = 7, nausea = 7, diarrhoea = 5, constipation = 2, abdominal pain = 1, gastrointestinal upset = 1, loose stools = 1, vomiting = 2.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0
- Actual outcome for Knee: Rectal bleeding at 15 days; Group 1: 3/162, Group 2: 0/155
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cardiac disorders at 15 days; Group 1: 0/161, Group 2: 4/155; Comments: No information on what the events were
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Nervous system disorders at 15 days; Group 1: 13/161, Group 2: 18/155; Comments: No information on what the events were
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 1, Reason: 1 person excluded (reason not given); Group 2 Number missing: 0 RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal disorders at 15 days; Group 1: 31/148, Group 2: 27/155; Comments: Provides what the adverse events were for some participants (but not all). Paracetamol: Dyspepsia = 11, nausea = 8, diarrhoea = 3, constipation = 1, abdominal pain = 1, gastrointestinal upset = 1, loose stools = 1, vomiting = 2, pharyngolarygeal pain = 2. Placebo: Dyspepsia = 7, nausea = 7, diarrhoea = 5, constipation = 2, abdominal pain = 1, gastrointestinal upset = 1, loose stools = 1, vomiting = 2. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Knee: Rectal bleeding at 15 days; Group 1: 0/148, Group 2: 0/155 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiac disorders at 15 days; Group 1: 1/148, Group 2: 4/155; Comments: No information on what the events were Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system disorders at 15 days; Group 1: 5/148, Group 2: 18/155; Comments: No information on what the events were Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, gender, race, pain and stiffness; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Goldstein 2007[85]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1045)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis diagnosis (no additional information)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People 18 years or older with an osteoarthritis diagnosis, an expectation of requiring ongoing NSAID use for at least 3 months, and a clinical indication for daily low-dose aspirin use for cardiovascular prophylaxis.
Exclusion criteria	Those who requiring anticoagulant agent or corticosteroid use; had a history of gastric or duodenal surgery other than a simple oversew of an ulcer during the past 1 year; had a known history of gastroduodenal ulcers; gastrointestinal bleed within the prior year; coexisting oesophageal disease; had a serum creatinine exceed 2.0 mg/dL; had a history of cancer within the prior 5 years (except basal cell carcinoma); had an anticipated need for surgery or other invasive procedures during the study period; had clinically abnormal laboratory, biochemical or haematological paramaeters; were pregnant or lactating or not currently using an acceptable form of birth control.
Recruitment/selection of patients	The study took place between July 22, 2003 and July 28, 2004.
Age, gender and ethnicity	Age - Mean (SD): 56.7 (11.2). Gender (M:F): 361:684. Ethnicity: White = 755, Black = 141, Hispanic = 110, Asian = 23, Other = 16.
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (General (no statement about the type)).
Extra comments	Severity: Not stated/unclear Duration of symptoms (mean [SD]): 114.6 (103.1) months
Indirectness of population	No indirectness

Interventions	(n=529) Intervention 1: Non-specific NSAIDs with gastroprotection (oral) - Non-specific NSAIDs with gastroprotection. Naproxen 500mg twice daily plus lansoprazole 30mg once daily. Duration 12 weeks. Concurrent medication/care: All received 81 or 325mg aspirin once daily.. Indirectness: No indirectness (n=516) Intervention 2: Specific COX-2 inhibitors (oral) - Celecoxib. Celecoxib 200mg once daily. Duration 12 weeks. Concurrent medication/care: All received 81 or 325mg aspirin once daily.. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by TAP Pharmaceutical Products Inc, Lake forest, IL.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NON-SPECIFIC NSAIDS WITH GASTROPROTECTION versus CELECOXIB

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulceration rate at 12 weeks; Group 1: 38/428, Group 2: 42/426
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index, race, low-dose aspirin use, number of erosions at baseline, negative H pylori status, duration of osteoarthritis, baseline joint pain severity, no prior NSAID use for 90 days, substance use.; Group 1 Number missing: 84, Reason: Naproxen and lansoprazole: 84 discontinued (35 adverse events, including 4 serious adverse events, 6 joint pain, 6 protocol violations, 5 lost to follow-up, 7 GI symptoms, 25 other reasons); Group 2 Number missing: 79, Reason: Celecoxib: 79 discontinued (33 adverse events, including 4 serious adverse events, 11 protocol violations, 5 joint pain, 6 GI symptoms, 1 GI complication, 7 lost to follow-up, 16 other reasons).

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Mild palpitations at 12 weeks; Group 1: 1/529, Group 2: 0/516
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index, race, low-dose aspirin use, number of erosions at baseline, negative H pylori status, duration of osteoarthritis, baseline joint pain severity, no prior NSAID use for 90 days, substance use.; Group 1 Number missing: 84, Reason: Naproxen and lansoprazole: 84 discontinued (35 adverse events, including 4 serious adverse events, 6 joint pain, 6 protocol violations, 5 lost to follow-up, 7 GI symptoms, 25 other reasons); Group 2 Number missing: 79, Reason: Celecoxib: 79 discontinued (33 adverse events, including 4 serious adverse events, 11 protocol violations, 5 joint pain, 6 GI symptoms, 1 GI complication, 7 lost to follow-up, 16 other reasons).

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Gordo 2017[86]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=388)
Countries and setting	Conducted in Germany, Spain, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of osteoarthritis of the knee according to the American College of Rheumatology guidelines, in a flare state and with a Functional Capacity Class of I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People at least 40 years of age with a clinical diagnosis of osteoarthritis of the knee according to the American College of Rheumatology guidelines, in a flare state and with a Functional Capacity Class of I-III. For people who had been receiving NSAID or analgesic therapy, which was discontinued 48 hours prior to the baseline visit, osteoarthritis flare was defined as Patient's Assessment of Arthritis pain between 40 and 90mm on an 100mm visual analogue scale, and an increase of one or more grades in both the Patient's and Physician's Global Assessments of Arthritis (scored from 1 'very good' to 5 'very poor' based on impact of symptoms) between screening and baseline visit. In people whose osteoarthritis was not controlled and who had not been receiving treatment for their osteoarthritis, osteoarthritis flare was defined as Patient's Assessment of Arthritis Pain between 40 and 90mm on an 100-mm visual analogue scale, and a rating of 'poor' or 'very poor' in both Patient's and Physician's Global Assessment of Arthritis.
Exclusion criteria	Inflammatory arthritis or gout/pseudogout or had experience an acute flare within the past 2 years (people with fibromyalgia were not excluded); previously had or anticipated a need for surgical or other invasive procedures on the joint with osteoarthritis during the study; had received oral corticosteroids within 4 weeks or paracetamol within 24 hours; malignancy or history of malignancy; active gastrointestinal disease; history of gastrointestinal perforations; obstructions or bleeding; cardiac, renal and/or hepatic disease; coagulation disorders or known hypersensitivity to COX-2 inhibitors, aspirin, NSAIDs or sulfonamide medications

Recruitment/selection of patients	People previously on NSAIDs had a washout period and had to fulfill flare criteria. People previously not on NSAIDs also had to fulfill flare criteria, which were adapted for this cohort.
Age, gender and ethnicity	Age - Mean (SD): 63.0 (6.7). Gender (M:F): 106:282. Ethnicity: White = 381, Black = 2, Hispanic = 0, Asian = 4, Other = 1
Further population details	1. Age: Mixed (Based on range, 40-89). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of osteoarthritis: 5.7 (5.9) years
Indirectness of population	No indirectness
Interventions	(n=309) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 800mg three times daily and placebo capsule once a day or Celecoxib 200mg orally once daily and placebo tablets three times daily. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The ibuprofen and celecoxib groups were combined due to class effect as agreed in the protocol (n=79) Intervention 2: Placebo. Placebo capsule once a day and placebo tablets three times a day. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This work was supported by Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Patient's Assessment of Arthritis pain (visual analogue scale) at 6 weeks; Group 1: mean -33.7 (SD 25); n=245, Group 2: mean -28.4 (SD 25.5); n=56; VAS 0-100 Top=High is poor outcome; Comments: Reports mean change score and standard error. Converted to standard deviation. Reported celecoxib: -34.5 (2.23). Reported ibuprofen: -32.8 (2.28). Reported placebo: -28.4 (3.41). Calculated SD celecoxib: 24.6. Calculated SD ibuprofen: 25.3. Calculated SD placebo: 25.5.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 53, Reason: Ibuprofen: 11 adverse events, 1 lab data, 10 other, 5 defaulted. Celecoxib: 5 adverse events, 16 other, 5 defaulted; Group 2 Number missing: 23, Reason: Placebo: 5 adverse events, 0 lab data, 14 other, 4 defaulted.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Upper GI events (abdominal pain, dyspepsia and/or nausea) at 6 weeks; Group 1: 10/309, Group 2: 2/79; Comments: Ibuprofen: 8; Celecoxib: 2; Placebo: 2 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 53, Reason: Ibuprofen: 11 adverse events, 1 lab data, 10 other, 5 defaulted. Celecoxib: 5 adverse events, 16 other, 5 defaulted; Group 2 Number missing: 23, Reason: Placebo: 5 adverse events, 0 lab data, 14 other, 4 defaulted. Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 1/309, Group 2: 2/79 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 53, Reason: Ibuprofen: 11 adverse events, 1 lab data, 10 other, 5 defaulted. Celecoxib: 5 adverse events, 16 other, 5 defaulted; Group 2 Number missing: 23, Reason: Placebo: 5 adverse events, 0 lab data, 14 other, 4 defaulted.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Gottesdiener 2002[87]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=617)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (followed up by an additional 8 week trial where the randomisation was redone which may have affected results and so will not be included in the analysis)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical and radiographic evidence of osteoarthritis of the tibiofemoral joint of the knee. Radiographic criteria were joint-space narrowing with the presence of osteophytes.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were a minimum of 40 years old and had clinical and radiographic evidence of osteoarthritis of the tibiofemoral joint of the knee. The study joint had to be the primary source of pain or diability. People were in American Rheumatism Association functional class I-III. All people required NSAIDs for their osteoarthritis pain for at least 25 of the 30 days prior to screening. People who satisfied the entry criteria discontinued their prior NSAID therapy. Following a washout period of 3-15 days (depending on the dose and half-life of the prior therapy), these patients' Walking Pain (question 1 of the WOMAC pain subscale) was assessed on a patient-reported 100mm visual analogue scale. People were randomised to the study if they had moderate walking pain (at least 40mm on a visual analogue scale), a minimum increase (worsening) in walking pain (15mm visual analogue scale) and an increase (worsening) in the Investigator's Assessment of Disease Status of 1 point (on a 5 point Likert scale) compared to values obtained at screening while people were receiving their prior NSAID therapy.
Exclusion criteria	Significant renal impairment; clinically significant abnormalities on screening physical or laboratory examinations (calculated creatinine clearance of no more than 30mL/min); class III/IV angina or uncontrolled congestive heart failure; uncontrolled hypertension, stroke or a transient ischaemic attack within 2 years
Recruitment/selection of patients	People had to achieve a flare criteria after washout from NSAIDs
Age, gender and ethnicity	Age - Mean (range): 61.3 (40-87). Gender (M:F): 175:442. Ethnicity: Hispanic Americans = 22, Native American = 3, Black = 43, White = 549

Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: American Rheumatism Association functional class I-III (median class II) Duration of symptoms (mean): 7.83 years.
Indirectness of population	No indirectness
Interventions	(n=557) Intervention 1: NSAIDs - Etoricoxib. Etoricoxib 5, 10, 30, 60 or 90mg once daily. Duration 6 weeks. Concurrent medication/care: After completing 2 weeks of treatment, people were provided open-label paracetamol, maximum daily dose of 2.6g, that could be taken for osteoarthritic pain that was not adequately controlled by the study medication. Indirectness: No indirectness Comments: The groups with different doses of medication were combined due to class effect as agreed in the protocol. (n=60) Intervention 2: Placebo. Matching placebo once a day. Duration 6 weeks. Concurrent medication/care: After completing 2 weeks of treatment, people were provided open-label paracetamol, maximum daily dose of 2.6g, that could be taken for osteoarthritic pain that was not adequately controlled by the study medication. Indirectness: No indirectness
Funding	Study funded by industry (This research was funded by Merck Research Laboratories)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ETORICOXIB versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Upper GI perforation, ulceration or bleeding and lower GI bleeds at 6 weeks; Group 1: 1/557, Group 2: 0/60; Comments: Etoricoxib: 1 serious lower GI bleed
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, race, age, functional class, height, weight, osteoarthritis duration, and baseline values of outcomes; Group 1 Number missing: 57, Reason: Etoricoxib 5mg: adverse events = 2, lack of efficacy = 6, other reasons = 4. Etoricoxib 10mg: adverse events = 3, lack of efficacy = 11, other reasons = 3. Etoricoxib 30mg: adverse events = 3, lack of efficacy = 4, other reasons = 1. Etoricoxib 60mg: adverse events = 3, lack of efficacy = 2, other reasons = 4. Etoricoxib 90mg: adverse events = 6, lack of efficacy = 2, other reasons = 3. ; Group 2 Number missing: 10, Reason: adverse events = 2, lack of efficacy = 7, other reasons = 1.

Protocol outcome 2: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Acute renal failure, increased ALT/AST at 6 weeks; Group 1: 1/557, Group 2: 0/60; Comments: Acute renal failure = 0. Increased ALT/AST = 1 person in the 10mg etoricoxib group.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, race, age, functional class, height, weight, osteoarthritis duration, and baseline values of outcomes; Group 1 Number missing: 57, Reason: Etoricoxib 5mg: adverse events = 2, lack of efficacy = 6, other reasons = 4. Etoricoxib 10mg: adverse events = 3, lack of efficacy = 11, other reasons = 3. Etoricoxib 30mg: adverse events = 3, lack of efficacy = 4, other reasons = 1. Etoricoxib 60mg: adverse events = 3, lack of efficacy = 2, other reasons = 4. Etoricoxib 90mg: adverse events = 6, lack of efficacy = 2, other reasons = 3. ; Group 2 Number missing: 10, Reason: adverse events = 2, lack of efficacy = 7, other reasons = 1. Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 10/557, Group 2: 2/60; Comments: Etoricoxib 5mg = 3, etoricoxib 10mg = 3, etoricoxib 30mg = 2, etoricoxib 60mg = 0, etoricoxib 90mg = 2, placebo = 2 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, race, age, functional class, height, weight, osteoarthritis duration, and baseline values of outcomes; Group 1 Number missing: 57, Reason: Etoricoxib 5mg: adverse events = 2, lack of efficacy = 6, other reasons = 4. Etoricoxib 10mg: adverse events = 3, lack of efficacy = 11, other reasons = 3. Etoricoxib 30mg: adverse events = 3, lack of efficacy = 4, other reasons = 1. Etoricoxib 60mg: adverse events = 3, lack of efficacy = 2, other reasons = 4. Etoricoxib 90mg: adverse events = 6, lack of efficacy = 2, other reasons = 3. ; Group 2 Number missing: 10, Reason: adverse events = 2, lack of efficacy = 7, other reasons = 1.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months

Study	Grace 1999[88]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=74)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic and radiologic osteoarthritis of the knee requiring daily drug therapy, disease duration of at least 3 months
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Age >35 years, symptomatic and radiologic osteoarthritis of the knee requiring daily drug therapy, disease duration of at least 3 months, and clinical laboratory, haematological, biochemical and urinalysis values within ±10% of the normal range
Exclusion criteria	Stage 4 osteoarthritis; recent or current alcohol abuse; drug dependency; serious psychological disease; women who were pregnant, lactating, or of child-bearing potential not using an effective form of birth control; subjects with a significant history of allergies, corticosteroid, or hyaluronic acid injections of target knee within one month prior to enrollment; hypersensitivity to an NSAID; local skin disease; prior joint replacement surgery on the target knee; people who had started physiotherapy in the preceding 2 weeks or had anticipating stopping or starting physiotherapy during the trial; those who had donated blood in the previous 56 days or undergone multiple blood sampling 30 days before study outset; those who possessed a language or psychological barrier; intercurrent illness, adverse event or surgery; symptoms or signs indicating possible toxicity.
Recruitment/selection of patients	During the screening visit, eligible people were instructed to stop their current NSAID therapy between 3 and 7 days prior to the enrollment visit. People were given a supply of 500mg paracetamol tablets during the screening visit with a dose regimen of 2 tablets per dose for control of pain not to exceed 8 tablets per day. No other concomitant medications were allowed. If a person met the necessary flare criteria (persistent symptoms of osteoarthritis requiring daily use of medication) during the washout period, final enrollment was confirmed
Age, gender and ethnicity	Age - Mean (SD): 62.0 (13.0). Gender (M:F): 29:45. Ethnicity: Caucasian = 72, Other = 2

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Low comorbidity score (People with chronic comorbidity: 37, people without: 35.). 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: ACR grade 1-4, median grade 3 Duration of symptoms: 135.8 (163) months.
Indirectness of population	No indirectness
Interventions	(n=38) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac gel 2.5g rubbed into the affected area with 2 fingers for between 5 and 20s.. Duration 2 weeks. Concurrent medication/care: People were allowed to maintain normal physical activities. Paracetamol was allowed as rescue medications (1000mg up to three times daily).. Indirectness: No indirectness (n=36) Intervention 2: Placebo. Matching placebo gel. Duration 2 weeks. Concurrent medication/care: People were allowed to maintain normal physical activities. Paracetamol was allowed as rescue medications (1000mg up to three times daily).. Indirectness: No indirectness
Funding	Study funded by industry (Study conducted by the University of Alberta and J.A.R Pharmaceuticals Ltd., Edmonton, Alberta, Canada)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 2 weeks; Group 1: mean -16.49 (SD 15.16); n=34, Group 2: mean -4.35 (SD 22.55); n=34; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 44.68 (17.50). Baseline placebo: 39.77 (17.89).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, height, weight, target knee, symptomatic involvement, chronic co-morbidity, acute intermittent illness, ACR grade, duration of osteoarthritis, and baseline values for WOMAC scores; Group 1 Number missing: 4, Reason: 1 withdrew after 5 days due to a rash. Other people it is unclear why they were not included.; Group 2 Number missing: 2, Reason: 2 had not completed all visits, 1 terminated due to protocol violation. However, this is one additional person than who appears to be excluded. Unclear.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 2 weeks; Group 1: mean -11.96 (SD 13.37); n=34, Group 2: mean -3.17 (SD 17.72); n=34; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 46.37 (16.51). Baseline placebo: 40.61 (18.14).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, height, weight, target knee,

symptomatic involvement, chronic co-morbidity, acute intermittent illness, ACR grade, duration of osteoarthritis, and baseline values for WOMAC scores; Group 1 Number missing: 4, Reason: 1 withdrew after 5 days due to a rash. Other people it is unclear why they were not included.; Group 2 Number missing: 2, Reason: 2 had not completed all visits, 1 terminated due to protocol violation. However, this is one additional person than who appears to be excluded. Unclear. Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Nausea and cramps at 2 weeks; Group 1: 1/38, Group 2: 2/36; Comments: Diclofenac: 1 = nausea and cramps. Placebo: 2 = nausea Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, height, weight, target knee, symptomatic involvement, chronic co-morbidity, acute intermittent illness, ACR grade, duration of osteoarthritis, and baseline values for WOMAC scores; Group 1 Number missing: 4, Reason: 1 withdrew after 5 days due to a rash. Other people it is unclear why they were not included.; Group 2 Number missing: 2, Reason: 2 had not completed all visits, 1 terminated due to protocol violation. However, this is one additional person than who appears to be excluded. Unclear. Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Numbness at 2 weeks; Group 1: 0/38, Group 2: 1/36 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, height, weight, target knee, symptomatic involvement, chronic co-morbidity, acute intermittent illness, ACR grade, duration of osteoarthritis, and baseline values for WOMAC scores; Group 1 Number missing: 4, Reason: 1 withdrew after 5 days due to a rash. Other people it is unclear why they were not included.; Group 2 Number missing: 2, Reason: 2 had not completed all visits, 1 terminated due to protocol violation. However, this is one additional person than who appears to be excluded. Unclear.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Haghighi 2005[89]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical dysfunction and pain due to osteoarthritis with radiological verification
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with clinical dysfunction and pain due to osteoarthritis with radiological verification with hip or knee pain on movement of >30mm on a 100mm visual analogue pain scale.
Exclusion criteria	Rheumatoid arthritis; metabolic disorders (diabetes); gastrointestinal disorders (gastritis or duodenum ulcer); neurological disorders; dementia
Recruitment/selection of patients	People taking previous medication underwent a washout of 1 week. Did not appear to have a baseline assessment after this period.
Age, gender and ethnicity	Age - Mean (range): 58.5 (52-64). Gender (M:F): 89:31. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 400mg three times daily. Duration 4 weeks. Concurrent medication/care: Paracetamol was used as a rescue medication throughout the study (1 to 3 tablets daily). Treated with analgesics and NSAIDs was discontinued during the one week washout period.. Indirectness: No indirectness (n=40) Intervention 2: Placebo. Matching placebo. Duration 4 weeks. Concurrent medication/care: Paracetamol was used as a rescue medication throughout the study (1 to 3 tablets daily). Treated with analgesics and NSAIDs was discontinued during the

	one week washout period.. Indirectness: No indirectness (n=40) Intervention 3: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). 30mg ginger extract in two 500mg capsules. Duration 4 weeks. Concurrent medication/care: Paracetamol was used as a rescue medication throughout the study (1 to 3 tablets daily). Treated with analgesics and NSAIDs was discontinued during the one week washout period.. Indirectness: No indirectness Comments: Ginger extract is not an intervention included in the protocol so is not included in the analysis. It is reported here for completeness.
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Visual analogue scale at 4 weeks; Group 1: mean 28 (SD 21.5); n=40, Group 2: mean 56.5 (SD 22.8); n=40; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports final values and standard error. Converted to standard deviation. Reported ibuprofen: 28 (3.4). Reported placebo: 56.5 (3.6). Baseline ibuprofen: 71.2 (2.4). Baseline placebo: 64.2 (2.8). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, range, sex and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Herrero-beaumont 2007[90]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=318)
Countries and setting	Conducted in Portugal, Spain; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary symptomatic knee osteoarthritis (in 1 or both knees) according to the clinical and radiographic criteria of the American College of Rheumatology
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients who were seen in rheumatology clinics at the participating centers if diagnosed as having primary symptomatic knee osteoarthritis (in 1 or both knees) according to the clinical and radiographic criteria of the American College of Rheumatology. Enrollment of people who were obese (BMI >30 kg/m²) was discouraged by the study protocol, to avoid any bias introduced by this. Disease stage was determined based on the Kellgren and Lawrence radiographic system and symptom severity was quantified with algofunctional indexes selected as outcome measures.
Exclusion criteria	Standard exclusion criteria (not defined)
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 63.9 (7.0). Gender (M:F): 40:278. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms (mean [SD]): 7.0 (5.7) years
Indirectness of population	No indirectness
Interventions	(n=108) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 1 gram three times a day and placebo to match glucosamine and matching placebo. Duration 26 weeks. Concurrent medication/care: The rescue medication consisted of the conventional nonsteroidal antiinflammatory drug ibuprofen, in 400mg tablets. People were instructed to : 1) leave the painful joint to rest for at least 1 hour, 2) take 1 ibuprofen tablet every 8 hours, 3) limit intake to a maximum of 3 days, 4) if needed,

	resume the rescue medication intake after a washout period of at least 7 days, 5) in all cases, suspend any use of rescue medication at least 7 days before a clinic visit. Indirectness: No indirectness (n=106) Intervention 2: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). 1 sachet of 1500mg glucosamine sulfate oral powder for solution per day and matching placebo. Duration 26 weeks. Concurrent medication/care: The rescue medication consisted of the conventional nonsteroidal antiinflammatory drug ibuprofen, in 400mg tablets. People were instructed to : 1) leave the painful joint to rest for at least 1 hour, 2) take 1 ibuprofen tablet every 8 hours, 3) limit intake to a maximum of 3 days, 4) if needed, resume the rescue medication intake after a washout period of at least 7 days, 5) in all cases, suspend any use of rescue medication at least 7 days before a clinic visit. Indirectness: No indirectness (n=104) Intervention 3: Placebo. Matching placebo (tablets three times a day and a sachet once a day). Duration 26 weeks. Concurrent medication/care: The rescue medication consisted of the conventional nonsteroidal antiinflammatory drug ibuprofen, in 400mg tablets. People were instructed to : 1) leave the painful joint to rest for at least 1 hour, 2) take 1 ibuprofen tablet every 8 hours, 3) limit intake to a maximum of 3 days, 4) if needed, resume the rescue medication intake after a washout period of at least 7 days, 5) in all cases, suspend any use of rescue medication at least 7 days before a clinic visit. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Rottapharm)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus GLUCOSAMINE (LICENSED PREPARATIONS ONLY)

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 26 weeks; Group 1: mean -2.4 (SD 3.2); n=108, Group 2: mean -2.7 (SD 3.2); n=106; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported paracetamol: -2.4 (-3.0 to -1.8). Reported glucosamine: -2.7 (-3.3 to -2.1). Calculated SD paracetamol: 3.2. Calculated SD glucosamine: 3.2. Baseline paracetamol: 8.0 (2.9). Baseline glucosamine: 7.8 (3.0).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 26 weeks; Group 1: mean -8.7 (SD 10.1); n=108, Group 2: mean -9.2 (SD 10.5); n=106; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported paracetamol: -8.7 (-10.6 to -6.8). Reported glucosamine: -9.2 (-11.2 to -7.2). Calculated SD paracetamol: 10.1. Calculated SD glucosamine: 10.5. Baseline paracetamol: 29.4 (11.0). Baseline glucosamine: 27.8 (11.4).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Precordial chest pain, atrial flutter at 26 weeks; Group 1: 1/108, Group 2: 0/106; Comments: Paracetamol: Atrial flutter = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.

Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Clinically significant abnormalities in transaminases and/or gamma glutamyl transferase (GGT) at 26 weeks; Group 1: 21/108, Group 2: 2/106; Comments: Abnormalities in liver function (transaminases and gamma glutamyl transferase)...were detected in 21 people in the paracetamol group versus 2 and 6 in the glucosamine and placebo groups respectively
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.violations.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 26 weeks; Group 1: mean -2.4 (SD 3.2); n=108, Group 2: mean -1.8 (SD 3.9); n=104; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported paracetamol: -2.4 (-3.0 to -1.8). Reported placebo: -1.8 (-2.6 to -1.1). Calculated SD paracetamol: 3.2. Calculated SD placebo: 3.9. Baseline paracetamol: 8.0 (2.9). Baseline placebo: 7.9 (3.0).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 26 weeks; Group 1: mean -8.7 (SD 10.1); n=108, Group 2: mean -5.5 (SD 11.5); n=104; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported paracetamol: -8.7 (-10.6 to -6.8). Reported placebo: -5.5 (-7.7 to -3.3). Calculated SD paracetamol: 10.1. Calculated SD placebo: 11.5. Baseline paracetamol: 29.4 (11.0). Baseline placebo: 27.2 (10.9).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Precordial chest pain, atrial flutter at 26 weeks; Group 1: 1/108, Group 2: 1/104; Comments: Paracetamol: Atrial flutter = 1. Placebo: Precordial chest pain = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations
Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Clinically significant abnormalities in transaminases and/or gamma glutamyl transferase (GGT) at 26 weeks; Group 1: 21/108, Group 2: 6/104; Comments: Abnormalities in liver function (transaminases and gamma glutamyl transferase)...were detected in 21 people in the paracetamol group versus 2 and 6 in the glucosamine and placebo groups respectively
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Paracetamol: 12 adverse events, 5 lack of efficacy, 3 lost to follow up, 8 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 26 weeks; Group 1: mean -2.7 (SD 3.2); n=106, Group 2: mean -1.8 (SD 3.9); n=104; WOMAC pain

subscale 0-20 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported glucosamine: -2.7 (-3.3 to -2.1). Reported placebo: -1.8 (-2.6 to -1.1). Calculated SD glucosamine: 3.2. Calculated SD placebo: 3.9. Baseline glucosamine: 7.8 (3.0). Baseline placebo: 7.9 (3.0).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 26 weeks; Group 1: mean -9.2 (SD 10.5); n=106, Group 2: mean -5.5 (SD 11.5); n=104; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean change scores and 95% confidence intervals. Converted into standard deviation. Reported glucosamine: -9.2 (-11.2 to -7.2). Reported placebo: -5.5 (-7.7 to -3.3). Calculated SD glucosamine: 10.5. Calculated SD placebo: 11.5. Baseline glucosamine: 27.8 (11.4). Baseline placebo: 27.2 (10.9).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Precordial chest pain, atrial flutter at 26 weeks; Group 1: 0/106, Group 2: 1/104; Comments: Placebo: Precordial chest pain = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Clinically significant abnormalities in transaminases and/or gamma glutamyl transferase (GGT) at 26 weeks; Group 1: 2/106, Group 2: 6/105; Comments: Abnormalities in liver function (transaminases and gamma glutamyl transferase)...were detected in 21 people in the paracetamol group versus 2 and 6 in the glucosamine and placebo groups respectively
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of knee osteoarthritis, mild knee joint effusion, Kellgren Lawrence grade, Lequesne index, and WOMAC scores; Group 1 Number missing: 28, Reason: Glucosamine: 4 adverse events, 7 lack of efficacy, 5 lost to follow up, 12 protocol violations.; Group 2 Number missing: 34, Reason: Placebo: 9 adverse events, 8 lack of efficacy, 5 lost to follow up, 12 protocol violations

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Houpt 1999[93]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=118)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	2nd line
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate pain due to primary knee osteoarthritis for at least 6 months. All people had to have radiological changes.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with at least "moderate" pain due to primary knee osteoarthritis for at least 6 months. All people had to have radiological changes (joint space narrowing and/or osteophytes), had never taken glucosamine products or oral corticosteroids, and had no recent (within 6 months) intraarticular injections, between the ages of 40 and 85 years and able to walk without assistive devices other than a cane
Exclusion criteria	No additional criteria
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 64.5 (9.8). Gender (M:F): 45:73. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Radiological score median grade 2. Includes people with grades 1-4. Duration of symptoms (mean [SD]): 8.3 (8.3) years. No statement that glucosamine was quality assured
Indirectness of population	No indirectness
Interventions	(n=58) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine hydrochloride 500mg three times a day. Duration 8 weeks. Concurrent medication/care: People were asked not to attend physiotherapy or take nonsteroidal antiinflammatory drugs or analgesics or use topical or intraarticular agents. People were permitted to use paracetamol 500mg up to a maximum of eight capsules per day.. Indirectness: No indirectness (n=60) Intervention 2: Placebo. Placebo three times a day. Duration 8 weeks.

	Concurrent medication/care: People were asked not to attend physiotherapy or take nonsteroidal antiinflammatory drugs or analgesics or use topical or intraarticular agents. People were permitted to use paracetamol 500mg up to a maximum of eight capsules per day.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by a grant-in-aid of research from Wampole Canada Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 8 weeks; Group 1: mean 7.14 (SD 4.01); n=58, Group 2: mean 7.65 (SD 4.13); n=60; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline glucosamine: 8.84 (2.95). Baseline placebo: 8.40 (3.59).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, years of symptoms, only NSAID use prior to study, only analgesics prior to study, prior use of both NSAIDs and analgesics, WOMAC, radiological score and pain at baseline; Group 1 Number missing: 13, Reason: 13 people discontinued (reason not given); Group 2 Number missing: 7, Reason: 7 people discontinued (reason not given)

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 8 weeks; Group 1: mean 25.98 (SD 14.7); n=58, Group 2: mean 27.17 (SD 14.1); n=60; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline glucosamine: 33.43 (9.76). Baseline placebo: 30.13 (10.8).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, years of symptoms, only NSAID use prior to study, only analgesics prior to study, prior use of both NSAIDs and analgesics, WOMAC, radiological score and pain at baseline; Group 1 Number missing: 13, Reason: 13 people discontinued (reason not given); Group 2 Number missing: 7, Reason: 7 people discontinued (reason not given)

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Mild gastrointestinal symptoms, such as gas, abdominal bloating and/or cramps at 8 weeks; Group 1: 7/58, Group 2: 7/60; Comments: "About 12% of both the glucosamine and placebo group had mild gastrointestinal symptoms".
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, weight, years of symptoms, only NSAID use prior to study, only analgesics prior to study, prior use of both NSAIDs and analgesics, WOMAC, radiological score and pain at baseline; Group 1 Number missing: 13, Reason: 13 people discontinued (reason not given); Group 2 Number missing: 7, Reason: 7 people discontinued (reason not given)

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Hubault 1976[94]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=9)
Countries and setting	Conducted in Unknown; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention time: 4 weeks (2 weeks for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with painful osteoarthritis of the hip made on the basis of the usual clinical criteria, together with positive radiological signs
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People with painful osteoarthritis of the hip who had already responded favourably to non-steroidal antiinflammatory treatment during the previous year
Exclusion criteria	People with osteoarthritis of the hip showing rapid destruction of the joint; those recently treated by surgery; pregnant women; people with a recognised contraindication to anti-inflammatory treatment such as a history of gastroduodenal ulcer or an existing severe hepatic, renal or haematological disorder
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Other: Not stated. Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: Not stated / Unclear 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=9) Intervention 1: NSAIDs - Ketoprofen. Ketoprofen 50mg three times a day. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=9) Intervention 2: Placebo. Placebo three times a day. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hip: Heartburn, nausea, diarrhoea at 2 weeks; Group 1: 2/9, Group 2: 2/9; Comments: Ketoprofen: 1 case of heartburn, 1 case of diarrhoea. Placebo: 1 case of nausea, 1 case of diarrhoea (same person as the ketoprofen arm).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Hip: Increase in urinary red and white blood cells at 2 weeks; Group 1: 1/9, Group 2: 0/9
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	**Hudson 2021[95]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=205)
Countries and setting	Conducted in New Zealand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis assessed by an orthopaedic surgeon through referral letters and accompanying pre-referral X-rays
Stratum	Knee
Subgroup analysis within study	Not applicable

Inclusion criteria	Have primary knee osteoarthritis defined according to the American College of Rheumatology clinical criteria for the classification of idiopathic osteoarthritis of the knee; have knee pain scoring at least 2 points on the WOMAC pain subscale (0-50 range); and have been on a stable analgesic regime for at least the previous 2 months.
Exclusion criteria	Previous joint replacement of the study knee; intra-articular steroid injection in the previous 3 months; secondary osteoarthritis; concurrent use of any antidepressant; any established contraindications to tricyclic antidepressants.
Recruitment/selection of patients	People were largely recruited from urban and suburban areas of Christchurch, New Zealand's second largest city. Invitation letters to participate were sent to people with knee osteoarthritis who had been declined specialist orthopaedic assessment for knee replacement (by a referral triaging orthopaedic surgeon, who had read the referral letters and examined accompanying pre-referral X-rays) and had returned to their GP for ongoing care.
Age, gender and ethnicity	Age - Mean (SD): 64.5 (9.2). Gender (M:F): 118:87. Ethnicity: European = 183, Maori = 21, Other = 16.
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: High comorbidity score (People with chronic conditions = 121). 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated/unclear Duration of symptoms (mean [SD]): 7.6 (7.6) years
Indirectness of population	No indirectness
Interventions	(n=102) Intervention 1: Antidepressants (oral) - Tricyclics. Nortryptyline 25mg daily for 2 weeks, after which the dose could be adjusted (over the next 6 weeks in 2 week intervals) to a maximum dose of four capsules daily (100mg nortriptyline). At 8 weeks, people were instructed to maintain their current dose until week 14.. Duration 14 weeks. Concurrent medication/care: People were free to use and adjust their usual analgesic medication as prescribed by their GP, but were requested not to use any other antidepressants or receive intra-articular steroid injections.. Indirectness: No indirectness (n=103) Intervention 2: Placebo. Matching placebo. Duration 14 weeks. Concurrent medication/care: People were free to use and adjust their usual analgesic medication as prescribed by their GP, but were requested not to use any other antidepressants or receive intra-articular steroid injections.. Indirectness: No indirectness
Funding	Academic or government funding (Funded with a project grant from the Health Research Council of New Zealand (reference number: 14/152).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRICYCLICS versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: RAND-36 (SF-36) physical function at 14 weeks; Group 1: mean 2 (SD 10); n=102, Group 2: mean -0.6 (SD 8.8); n=103; SF-36 physical function 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 29.9 (10.6). Baseline placebo: 31.6 (9.9).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed

at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) role physical at 14 weeks; Group 1: mean 1.8 (SD 11.9); n=102, Group 2: mean -0.1 (SD 11.5); n=103; SF-36 role physical 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 39.9 (11.0). Baseline placebo: 31.6 (9.9).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) bodily pain at 14 weeks; Group 1: mean 5.8 (SD 8.8); n=102, Group 2: mean 3.1 (SD 9.4); n=103; SF-36 bodily pain 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 35.7 (8.1). Baseline placebo: 35.5 (7.5).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) general health at 14 weeks; Group 1: mean 0.1 (SD 7.3); n=102, Group 2: mean 0.6 (SD 7.8); n=103; SF-36 general health 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 47.4 (8.4). Baseline placebo: 46.1 (8.6).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) vitality at 14 weeks; Group 1: mean 0.6 (SD 8.5); n=102, Group 2: mean 0 (SD 9.9); n=103; SF-36 vitality 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 46.2 (8.8). Baseline placebo: 47.1 (10.0).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) social function at 14 weeks; Group 1: mean 0.4 (SD 13.2); n=102, Group 2: mean -1.6 (SD 12.8); n=103; SF-36 social function 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 44.7 (11.5). Baseline placebo: 44.8 (11.4).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) role emotional at 14 weeks; Group 1: mean -1.3 (SD 11.8); n=102, Group 2: mean -3.1 (SD 13.9); n=103; SF-36 role emotional 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 47.3 (11.8). Baseline placebo: 47.8 (11.4).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent - Actual outcome for Knee: RAND-36 (SF-36) mental health at 14 weeks; Group 1: mean -0.6 (SD 9.7); n=102, Group 2: mean -0.4 (SD 8.9); n=103; SF-36

mental health 0-100 Top=High is good outcome; Comments: Baseline nortriptyline: 51.9 (9.1). Baseline placebo: 52.1 (8.8).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain at 14 weeks; Group 1: mean -24.3 (SD 22.5); n=102, Group 2: mean -18.7 (SD 25.8); n=103; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline nortriptyline: 60.2 (13.5). Baseline placebo: 61.2 (12.5).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function at 14 weeks; Group 1: mean -23.2 (SD 21.5); n=102, Group 2: mean -18 (SD 23.2); n=103; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline nortriptyline: 62.8 (15.0). Baseline placebo: 59.9 (14.8).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Hospitalisation for atrial fibrillation or life-threatening myocardial infarction at 14 weeks; Group 1: 2/99, Group 2: 0/102; Comments: Tricyclics: 1 life-threatening myocardial infarction, 1 hospitalisation for atrial fibrillation
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent

Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Hospitalisation for renal calculi at 14 weeks; Group 1: 0/99, Group 2: 1/102
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent

Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

- Actual outcome for Knee: Headache at 14 weeks; Group 1: 14/99, Group 2: 27/102 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, ethnicity, BMI, years with knee ostoearthritis, use of assistive devices, chronic conditions and baseline values of outcomes; Group 1 Number missing: 3, Reason: 99 people were assessed at week 14 - 1 withdrew consent, 2 lost to follow up; Group 2 Number missing: 1, Reason: 102 people were assessed at week 14 - 1 withdrew consent	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

Study	Hughes 2002[96]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 weeks (6 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with radiologically defined, symptomatic osteoarthritis of at least one knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with radiologically defined, symptomatic osteoarthritis of at least one knee; pain or discomfort should have been experienced in the affected knee on most days for the previous 3 months. People who had an arthroplasty of one knee were not excluded as long as the other knee was affected and symptomatic.
Exclusion criteria	People if they had prosthetic material in both knees; had previously taken glucosamine; had received an arthroscopic washout of the index knee in the previous 3 months or had any intra-articular injection to the knee within the previous month
Recruitment/selection of patients	Pragmatic trial designed to include a wide range of pain severity and all grades of radiological severity of osteoarthritis of the knee.
Age, gender and ethnicity	Age - Mean (SD): 62.28 (9.12). Gender (M:F): 26:54. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren and Lawrence grades 1-4 - median grade 3. Duration of osteoarthritis (mean [SD]): 7.62 (8.06) years.
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulphate 1500mg a day taken as 500mg capsules three times a day/ The capsules contained: 500mg of potassium chloride-free glucosamine sulphate, 300mg of vitamin C, 300mg calcium carbonate and 5mg of manganese.. Duration 24 weeks. Concurrent medication/care: All people were permitted to continue their existing NSAID medication and encouraged to avoid changing their dose or medication during the course of the study. Use of NSAIDs was recorded at each follow up assessment. People were allowed access to paracetamol

	or other proprietary or prescribed simple analgesic and reported use of analgesia was collected using a patient diary and recorded at each follow-up assessment.. Indirectness: No indirectness (n=40) Intervention 2: Placebo. Matching placebo three times a day. Duration 24 weeks. Concurrent medication/care: All people were permitted to continue their existing NSAID medication and encouraged to avoid changing their dose or medication during the course of the study. Use of NSAIDs was recorded at each follow up assessment. People were allowed access to paracetamol or other proprietary or prescribed simple analgesic and reported use of analgesia was collected using a patient diary and recorded at each follow-up assessment.. Indirectness: No indirectness
Funding	Study funded by industry (We thank Health Perception UK for providing all the trial medication and packaging, for shipping the medication to the pharmacy, and for generating and managing the randomisation list. The research nurses for this trial were funded by an educational grant from Health Perception UK.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 24 weeks; Group 1: mean 7.7 (SD 4.1); n=39, Group 2: mean 7.5 (SD 2.9); n=39; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports area under the curve values (y axis = weeks). Values calculated by dividing reported values by the number of weeks (24). Reported glucosamine: 184.88 (98.79). Reported placebo: 179.32 (69.96). Baseline (all): 9.24 (3.50).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Downgraded for outcome reporting as estimating the results for WOMAC subscales and cannot be certain if the results are the true values; Indirectness of outcome: No indirectness ; Baseline details: Reports overall baseline values for the entire cohort (all pooled together). Cannot compare baselines.; Group 1 Number missing: 2, Reason: 1 withdrew consent before the trial started. 1 withdrew after the 1st assessment.; Group 2 Number missing: 3, Reason: 1 withdrew consent before the trial started. 2 withdrew after the 1st assessment.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 24 weeks; Group 1: mean 27.7 (SD 16.4); n=39, Group 2: mean 26.1 (SD 12.6); n=39; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Reports area under the curve values (y axis = weeks). Values calculated by dividing reported values by the number of weeks (24). Reported glucosamine: 665.1 (394.42). Reported placebo: 625.20 (301.92). Baseline (all): 32.90 (13.81).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Downgraded for outcome reporting as estimating the results for WOMAC subscales and cannot be certain if the results are the true values; Indirectness of outcome: No indirectness ; Baseline details: Reports overall baseline values for the entire cohort (all pooled together). Cannot compare baselines.; Group 1 Number missing: 2, Reason: 1 withdrew consent before the trial started. 1 withdrew after the

1st assessment.; Group 2 Number missing: 3, Reason: 1 withdrew consent before the trial started. 2 withdrew after the 1st assessment.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	IPSO trial: Boureau 2004[34]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=222)
Countries and setting	Conducted in France; Setting: Outpatient follow up from general practice
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with painful osteoarthritis of the lower limbs (femorotibial or femoropatellar osteoarthritis of the knee or osteoarthritis of the hip) diagnosed according to the clinical and radiological criteria (X-ray examination performed within the previous year) from the American College of Rheumatology classification
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either sex, ages 50-85 years, with chronic pain (score at least 50mm on a 100mm visual analogue scale) due to confirmed knee or hip osteoarthritis and required analgesic treatment for 2 weeks. People with painful osteoarthritis of the lower limbs (femorotibial or femoropatellar osteoarthritis of the knee or osteoarthritis of the hip) diagnosed according to the clinical and radiological criteria from the American College of Rheumatology classification.
Exclusion criteria	Any serious respiratory, hepatic, or renal failure; prosthesis of the affected joint or surgery of the affected joint within the previous 3 months or planned within the next 3 months; a known hypersensitivity to any NSAID or paracetamol; need for anti-inflammatory treatment; having a hydrarthrosis requiring a puncture; use of oral corticosteroids within the previous 8 days
Recruitment/selection of patients	Multicentre trial
Age, gender and ethnicity	Age - Mean (SD): 66.5 (9.3). Gender (M:F): 60:162. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range (includes people up to 85 year olds)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: 4.7 (5.5) years
Indirectness of population	No indirectness

Interventions	(n=111) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 200mg two capsules three times daily for 14 days. Duration 14 days. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=111) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg two capsules three times daily for 14 days. Duration 14 days. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (The study was funded by Boots Healthcare France)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 14 days; Group 1: mean -23 (SD 18); n=111, Group 2: mean -14.5 (SD 16); n=111; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 50.0 (13.5). Baseline paracetamol: 50.0 (12.5).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 14 days; Group 1: mean -19.8 (SD 17.3); n=111, Group 2: mean -12.8 (SD 15); n=111; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 48.2 (13.8). Baseline paracetamol: 48.7 (12.9).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Digestive system adverse events at 14 days; Group 1: 14/111, Group 2: 10/111; Comments: Taken from listed outcome. Specific adverse events: Ibuprofen: nausea = 7, abdominal pain = 3, dyspepsia = 6, diarrhoea = 1. Paracetamol: nausea = 5, abdominal pain = 6, dyspepsia = 1, constipation = 2, diarrhoea = 1.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Cardiovascular system adverse events at 14 days; Group 1: 2/111, Group 2: 2/111; Comments: Taken from listed outcome. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason. Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Other: Urogenital system adverse events at 14 days; Group 1: 1/111, Group 2: 0/111; Comments: Taken from listed outcome. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason. Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Nervous system adverse events at 14 days; Group 1: 6/111, Group 2: 7/111; Comments: Taken from listed outcome. Specific adverse events: Ibuprofen: dizziness = 3, headache = 1, somnolence = 2, migraine = 1, paraesthesia = 1. Paracetamol: dizziness = 4, headache = 2, somnolence = 1, migraine = 1, paraesthesia = 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, body mass index, disease duration, location of osteoarthritis, and baseline values of outcomes; Group 1 Number missing: 16, Reason: Missing data for VAS = 9, a second drug intake started before the end of the first 6 hours = 3, study treatment started before the first pain evaluation = 4, forbidden concomitant treatment started = 2. Some people had more than one reason.; Group 2 Number missing: 17, Reason: Missing data for VAS = 12, a second drug intake started before the end of the first 6 hours = 5, study treatment started before the first pain evaluation = 1. Some people had more than one reason.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Kageyama 1973[97]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=89)
Countries and setting	Conducted in Japan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks (2 weeks for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Pain and/or swelling and/or limitation of movement. All had x-ray changes
Stratum	Multisite
Subgroup analysis within study	Not applicable
Inclusion criteria	Symptomatic people with no significant complications
Exclusion criteria	People with recent gastrointestinal bleeding
Recruitment/selection of patients	Other anti-arthritic therapy, including intra-articular corticosteroids and non-steroidal anti-inflammatory drugs, were stopped at least 2 weeks before the study was commenced
Age, gender and ethnicity	Age - Other: Reports that 39 people were <59 years, and 50 were >60 years. Gender (M:F): 20:69. Ethnicity: Not stated
Further population details	1. Age: Not stated / Unclear 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee, hip, fingers or polyarticular. Majority knee (72/89).).
Extra comments	Severity: Moderate to severe pain (median was moderately severe) Duration of disease: 67 were <6 years, 22 were >7 years..
Indirectness of population	No indirectness
Interventions	(n=74) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg daily (by three times a day dosing). Duration 2 weeks. Concurrent medication/care: Other anti-arthritis agents were not allowed during the study. Any pre-established physical therapy program was permitted to continue but maintained in a constant fashion during the study. Indirectness: No indirectness (n=43) Intervention 2: Placebo. Matching placebo three times daily. Duration 2 weeks. Concurrent medication/care: Other anti-arthritis agents were not allowed during the study. Any pre-established physical therapy program was permitted to continue but maintained in a constant fashion during the study. Indirectness: No indirectness

	(n=61) Intervention 3: NSAIDs - Other. Aluminium flufenamate 750mg daily (given as three tablets per day). Duration 2 weeks. Concurrent medication/care: Other anti-arthritis agents were not allowed during the study. Any pre-established physical therapy program was permitted to continue but maintained in a constant fashion during the study. Indirectness: No indirectness Comments: Aluminium Flufenamate is not licensed for use in the UK so was not included in the analysis as agreed in the protocol. It is reported here for completeness.
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Upper abdominal discomfort, slight headache, nausea, anorexia, vomiting at 2 weeks; Group 1: 5/74, Group 2: 5/43; Comments: Naproxen: Upper abdominal discomfort = 3, nausea = 1, nausea and anorexia =1. Placebo: Slight nausea = 2, nausea and vomiting = 1, upper abdominal discomfort = 1, nausea and anorexia = 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, weight, duration of disease, severity of symptoms, clinical findings and x-ray changes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Oedema (face and legs) at 2 weeks; Group 1: 1/74, Group 2: 0/43 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, weight, duration of disease, severity of symptoms, clinical findings and x-ray changes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 2 weeks; Group 1: 1/74, Group 2: 0/43 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, weight, duration of disease, severity of symptoms, clinical findings and x-ray changes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Karakaya 1977[98]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Turkey; Setting: Inpatient or outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 30 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical and radiological diagnosis of osteoarthritis of the knee joint
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of any sex, hospitalised or outpatients, after clinical and radiological diagnosis of osteoarthritis of the knee joint
Exclusion criteria	People with infectious arthritis or with conditions that needed to be treated surgery
Recruitment/selection of patients	One site (Bursa University)
Age, gender and ethnicity	Age - Other: Unclear. Gender (M:F): Unclear. Ethnicity: Not stated
Further population details	1. Age: Not stated / Unclear 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not reported Duration of symptoms: Not reported
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: NSAIDs - Indomethacin. Indomethacin 25mg three times a day. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=20) Intervention 2: NSAIDs - Other. Proquazone 300mg three times a day. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: Proquazone is not licensed for use in the UK and so was not included in this analysis as agreed in the protocol. It was reported here for completeness. (n=20) Intervention 3: Placebo. Placebo three times a day. Duration 4 weeks.

	Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Equipment / drugs provided by industry (Proquazone supplied as Biarison by Sandoz/Wander, Basle, Switzerland)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: INDOMETHACIN versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastric pain at 4 weeks; Group 1: 2/14, Group 2: 1/5 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Stated that there were no statistical differences between the three groups regarding sex, age, weight, height and severity of the disease but did not report the values; Group 1 Number missing: 8, Reason: 6 terminated due to inefficacy, 2 terminated due to side effects; Group 2 Number missing: 16, Reason: 15 terminated due to inefficacy, terminated due to side effects = 1	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Karlsson 2009[99]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=134)
Countries and setting	Conducted in Sweden; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of osteoarthritis of the hip and/or knee, based on American College of Rheumatology and radiographic criteria
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged >18 years with a clinical diagnosis of osteoarthritis of the hip and/or knee, based on American College of Rheumatology and radiographic criteria. Analgesia in the primary osteoarthritis joint had to be suboptimal in the week before the baseline visit, as evidenced by a Box Scal 11 (BS-11) score ≥4. In addition, people had to experience inadequate pain relief while taking immediate-release paracetamol 4000mg/day during the screening week.
Exclusion criteria	People with OA pain with high-potency opioid analgesics (eg. morphine, fentanyl, oxycodone, methadone, hydromorphone, ketobemidone, buprenorphine) or with a usual dose of tramadol, codeine or dextropropoxyphene for >1 week in the past 3 months; if the required frequent analgesic therapy for other chronic conditions (e.g. migraine, gout, rheumatoid arthritis); if they were scheduled for surgery during the screening or treatment phase of the study; if they were abusing controlled substances or alcohol or, in the opinion of the investigator, demonstrated behaviours suggestive of addiction or substance abuse; if they received a diagnosis of cancer (except basal cell carcinoma); had cancer in the past 5 years (except treated basal cell carcinoma).
Recruitment/selection of patients	People had to experience inadequate pain relief while taking immediate-release paracetamol 4000mg/day during the screening week.
Age, gender and ethnicity	Age - Mean (SD): 64.3 (10.3). Gender (M:F): 58:76. Ethnicity: White = 133, Asian = 1
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Osteoarthritis of the hip and/or knee).

Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=69) Intervention 1: Opioids (topical - systemic) - Opioids (topical systemic). 7 day buprenorphine patches. The possible doses were 5, 10, 15, and 20 micrograms/hour with up to 2 patches being able to be worn at the same time.. Duration 12 weeks. Concurrent medication/care: Paracetamol could be used for rescue medication. Indirectness: No indirectness (n=66) Intervention 2: Strong opioids (oral) - Tramadol. Tramadol twice daily - the possible doses were 150, 200, 300, 400 mg/day.. Duration 12 weeks. Concurrent medication/care: Paracetamol could be used for rescue medication. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored and designed by Mundipharma AB, Goteborg, Sweden)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OPIOIDS (TOPICAL SYSTEMIC) versus TRAMADOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Box scale-11 (numeric rating scale) pain at 12 weeks; Group 1: mean 3.92 (SD 2.07); n=69, Group 2: mean 4.1 (SD 2.15); n=65; Numeric rating scale 0-10 Top=High is poor outcome; Comments: Baseline buprenorphine: 6.16 (1.35). Baseline tramadol: 6.21 (1.55).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race and baseline pain values; Group 1 Number missing: 14, Reason: 10 adverse events, 3 protocol violations, 1 other; Group 2 Number missing: 21, Reason: 19 adverse events, 1 withdrawal of consent, 1 lost to follow up

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Peripheral oedema at 12 weeks; Group 1: 4/69, Group 2: 0/65
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race and baseline pain values; Group 1 Number missing: 14, Reason: 10 adverse events, 3 protocol violations, 1 other; Group 2 Number missing: 21, Reason: 19 adverse events, 1 withdrawal of consent, 1 lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-

	months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Kivitz 2001[104]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=491)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the knee for ≥6 months as confirmed by a weightbearing radiograph
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People who had osteoarthritis of the knee for at least 6 months as confirmed by a weightbearing radiograph and at least moderate pain flare within 24 hours prior to administration of the study medication. Eligible pain flare was defined by a change of more than or equal to 2 points from the screening visit (on a 10 point visual analogue scale).
Exclusion criteria	Knee joint instability; osteoarthritis of the hip ipsilateral to the index knee; inflammatory arthritis other than osteoarthritis; loss of articular cartilage as evidenced by the obliteration of joint space on knee X-ray; nasal polyps; history of bronchospasm or angioedema induced by NSAIDs; history within the past years of duodenal or peptic perforation or gastrointestinal haemorrhage requiring surgery; NSAID-induced gastric ulcer within the past 60 days; inflammatory bowel disease; or clinically significant clinical laboratory values.
Recruitment/selection of patients	People were included if they had a pain flare after the screening visit. Eligible pain flare was defined by a change of more than or equal to 2 points from the screening visit (on a 10 point visual analogue scale).
Age, gender and ethnicity	Age - Mean (range): 59.8 (28-91). Gender (M:F): 146:345. Ethnicity: White = 385, Black = 56, Hispanic = 44, Other = 6
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness

Interventions	(n=82) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 400mg once daily. Duration 2 weeks. Concurrent medication/care: Topical analgesics were allowed during the course of the study, provided they were not applied to the knee chosen for pain assessments. Topical and inhaled corticosteroids could also be continued if used at the time of study initiation. Aspirin ≤325mg/day could be used for cardiovascular prophylaxis.. Indirectness: No indirectness (n=160) Intervention 2: Placebo. Placebo once daily. Duration 2 weeks. Concurrent medication/care: Topical analgesics were allowed during the course of the study, provided they were not applied to the knee chosen for pain assessments. Topical and inhaled corticosteroids could also be continued if used at the time of study initiation. Aspirin ≤325mg/day could be used for cardiovascular prophylaxis.. Indirectness: No indirectness (n=249) Intervention 3: NSAIDs - Other. Oxaprozin potassium 1200mg with PRN 600mg oxaprozin potassium, or oxaprozin potassium 1800mg once per day. Duration 2 weeks. Concurrent medication/care: Topical analgesics were allowed during the course of the study, provided they were not applied to the knee chosen for pain assessments. Topical and inhaled corticosteroids could also be continued if used at the time of study initiation. Aspirin ≤325mg/day could be used for cardiovascular prophylaxis.. Indirectness: No indirectness Comments: Oxaprozin is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness.
Funding	Study funded by industry (This clinical research study was underwritten by a research grant from G.D. Searle & Co., Skokie, Illinois, USA.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PLACEBO

Protocol outcome 1: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 2 weeks; Group 1: 2/82, Group 2: 13/159
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, race, gender and baseline pain; Group 1 Number missing: 0, Reason: Reports of the 491 enrolled, 360 completed the study. 69 had treatment failure, 42 had adverse events, 18 had poor compliance, 2 were lost to follow up; Group 2 Number missing: 1, Reason: Reports of the 491 enrolled, 360 completed the study. 69 had treatment failure, 42 had adverse events, 18 had poor compliance, 2 were lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Kivitz 2001[105]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1061)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fulfillment of the American College of Rheumatology clinical and radiographic criteria for the diagnosis of primary osteoarthritis of the hip
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	Adult outpatients who fulfilled the American College of Rheumatology clinical and radiographic criteria for a diagnosis of primary osteoarthritis of the hip, were in a functional class of I-III and had symptomatic osteoarthritis flare at the baseline visit. Symptomatic osteoarthritis was demonstrated by a worsening of the signs and symptoms of the disease following discontinuation of conventional NSAID or other analgesic medications, between the screening and baseline visit, and by other criteria for people not receiving treatment. Pre-menopausal women were included in the study if they were not pregnant, and if they were taking adequate contraceptive measures. People taking up to 325mg/day aspirin for non-arthritic reasons, or with a history of gastrointestinal tract bleeding, fibrositis or fibromyalgia were not excluded unless they met other exclusion criteria.
Exclusion criteria	Received intra-articular or soft-tissue injections of corticosteroids within 4 weeks of receiving the first dose of study medication; a known hypersensitivity to COX-2 inhibitors, sulfonamides, or NSAIDs; received any investigational medication within 30 days of the first dose of study medication; taken any NSAIDs or any analgesic within 48 hours of the baseline assessment; received piroxicam and/or oxaprozin within 4 days of the baseline assessment; active, concomitant gastrointestinal tract, renal, hepatic or coagulation disorders; malignancy (unless in remission for 5 years); if they had been diagnosed with or treated for oesophageal/gastroduodenal ulceration within 30 days of receiving the study drug; people diagnosed with inflammatory arthritis, gout, or acute joint trauma at the hip; anticipated need for surgery during the study period.

Recruitment/selection of patients	People had to show evidence of a flare of osteoarthritis activity after withdrawal of medication (or by different criteria if not receiving treatment)
Age, gender and ethnicity	Age - Mean (range): 62 (28-93). Gender (M:F): 360:701. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Functional class of I-III Duration of symptoms (mean): 7.3 years.
Indirectness of population	No indirectness
Interventions	(n=843) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day (with corresponding placebo to match celecoxib group) or Celecoxib 100mg, 200mg or 400mg/day with matching placebo (taking medication twice a day). Duration 12 weeks. Concurrent medication/care: People were permitted to take up to 325mg aspirin daily, and up to 2g paracetamol daily for up to 3 consecutive days, when absolutely necessary for non-arthritic conditions. However, the use of paracetamol was prohibited within 48 hours of an assessment of arthritis efficacy. The use of oral or injectable corticosteroids within 4 weeks of the first dose of study drug was not permitted. Use of other medications during the trial was permitted, but was documented on the Concurrent Medications Diary Card.. Indirectness: No indirectness Comments: Naproxen and celecoxib groups were combined as agreed in the protocol (n=218) Intervention 2: Placebo. Matching placebo twice a day. Duration 12 weeks. Concurrent medication/care: People were permitted to take up to 325mg aspirin daily, and up to 2g paracetamol daily for up to 3 consecutive days, when absolutely necessary for non-arthritic conditions. However, the use of paracetamol was prohibited within 48 hours of an assessment of arthritis efficacy. The use of oral or injectable corticosteroids within 4 weeks of the first dose of study drug was not permitted. Use of other medications during the trial was permitted, but was documented on the Concurrent Medications Diary Card.. Indirectness: No indirectness
Funding	Study funded by industry (The writing of this article was supported by the Pharmacia Corporation and Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hip: Diarrhoea, dyspepsia, nausea, abdominal pain, constipation, flatulence at 12 weeks; Group 1: 231/843, Group 2: 39/218; Comments:

Including ... Celecoxib 100mg: Diarrhoea = 8, dyspepsia = 8, nausea = 8, abdominal pain = 8, constipation = 1, flatulence = 5. Celecoxib 200mg: Diarrhoea = 13, dyspepsia = 18, nausea = 12, abdominal pain = 11, constipation = 5, flatulence = 5. Celecoxib 400mg: Diarrhoea = 21, dyspepsia = 21, nausea = 12, abdominal pain = 10, constipation = 4, flatulence = 2. Naproxen: Diarrhoea = 11, dyspepsia = 19, nausea = 12, abdominal pain = 19, constipation = 8, flatulence = 12. Placebo: Diarrhoea = 12, dyspepsia = 16, nausea = 5, abdominal pain = 6, constipation = 2, flatulence = 3. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, duration of disease, aspirin users, and baseline values for outcomes; Group 1 Number missing: 286, Reason: Naproxen: 1 lost to follow up, 1 preexisting violation, 7 protocol noncompliance, 51 treatment failure. Celecoxib 100mg: 4 lost to follow up, 2 preexisting violation, 6 protocol noncompliance, 76 treatment failure. Celecoxib 200mg: 0 lost to follow up, 0 preexisting violation, 8 protocol noncompliance, 61 treatment failure. Celecoxib 400mg: 2 lost to follow up, 3 preexisting violation, 9 protocol noncompliance, 55 treatment failure.; Group 2 Number missing: 123, Reason: Placebo: 2 lost to follow up, 3 preexisting violation, 5 protocol noncompliance, 113 treatment failure Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Hip: Peripheral oedema at 12 weeks; Group 1: 23/843, Group 2: 1/218; Comments: Celecoxib 100mg: 3, celecoxib 200mg: 3, celecoxib 400mg: 11, naproxen: 6, placebo: 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, duration of disease, aspirin users, and baseline values for outcomes; Group 1 Number missing: 286, Reason: Naproxen: 1 lost to follow up, 1 preexisting violation, 7 protocol noncompliance, 51 treatment failure. Celecoxib 100mg: 4 lost to follow up, 2 preexisting violation, 6 protocol noncompliance, 76 treatment failure. Celecoxib 200mg: 0 lost to follow up, 0 preexisting violation, 8 protocol noncompliance, 61 treatment failure. Celecoxib 400mg: 2 lost to follow up, 3 preexisting violation, 9 protocol noncompliance, 55 treatment failure.; Group 2 Number missing: 123, Reason: Placebo: 2 lost to follow up, 3 preexisting violation, 5 protocol noncompliance, 113 treatment failure Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hip: Headache, dizziness at 12 weeks; Group 1: 166/843, Group 2: 50/218; Comments: Including ... Celecoxib 100mg: Headache = 26, dizziness = 3. Celecoxib 200mg: Headache = 43, dizziness = 4. Celecoxib 400mg: Headache = 34, dizziness = 6. Naproxen: Headache = 24, dizziness = 2. Placebo: Headache = 42, dizziness = 5. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, duration of disease, aspirin users, and baseline values for outcomes; Group 1 Number missing: 286, Reason: Naproxen: 1 lost to follow up, 1 preexisting violation, 7 protocol noncompliance, 51 treatment failure. Celecoxib 100mg: 4 lost to follow up, 2 preexisting violation, 6 protocol noncompliance, 76 treatment failure. Celecoxib 200mg: 0 lost to follow up, 0 preexisting violation, 8 protocol noncompliance, 61 treatment failure. Celecoxib 400mg: 2 lost to follow up, 3 preexisting violation, 9 protocol noncompliance, 55 treatment failure.; Group 2 Number missing: 123, Reason: Placebo: 2 lost to follow up, 3 preexisting violation, 5 protocol noncompliance, 113 treatment failure
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months;

	Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Kivitz 2002[102]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1019)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate to severe osteoarthritis of the knee according to the modified criteria of the American College of Rheumatology
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Ambulatory adults who had been diagnosed with moderate to severe osteoarthritis of the knee according to the modified criteria of the American College of Rheumatology. People who had baseline scores of at least 40mm on the Patient's Assessment of Arthritis Pain-Visual Analog Scale and baseline categorical scores of poor to very poor on the Patient's and Physician's Global Assessments of Arthritis.
Exclusion criteria	People with inflammatory arthritis; gout; pseudogout; Paget disease; any chronic pain syndrome; osteoarthritis of the hip ipsilateral to the index knee; severe anserine bursitis; acute joint trauma; complete loss of articular cartilage on the index knee; active gastrointestinal disease; gastrointestinal tract ulceration 30 days before the trial; a significant bleeding disorder; a history of gastric or duodenal surgery; an oesophageal, gastric, pyloric channel, or duodenal ulcer or a score of at least 10 for oesophageal, gastric, or dudodenal erosions at the pretreatment endoscopy examination.
Recruitment/selection of patients	People were recruited from primary care and rheumatology specialty settings.
Age, gender and ethnicity	Age - Mean (SD): 59.8 (10.9). Gender (M:F): 360:659. Ethnicity: White = 794, Black = 118, Asian = 6, Hispanic = 92
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Moderate to severe Duration of symptoms: 9.1 (8.5) years.
Indirectness of population	No indirectness
Interventions	(n=205) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice daily. Duration 12 weeks. Concurrent medication/care: People discontinued their normal medications at

	the following specified times before the baseline endoscopy: NSAIDs (including full dose aspirin at a dosage ≥325mg/day) at 48 hours, corticosteroid injection or hyaluronic acid injections at 3 and 6 months respectively. The use of antiulcer drugs, including histamine-2 receptor antagonists, proton pump inhibitors, misoprostol, and sucralfate was discontinued at least 24 hours before the baseline endoscopy.. Indirectness: No indirectness (n=609) Intervention 2: NSAIDs - Other. Valdecoxib 5mg, 10mg or 20mg once a day with placebo. Duration 12 weeks. Concurrent medication/care: People discontinued their normal medications at the following specified times before the baseline endoscopy: NSAIDs (including full dose aspirin at a dosage ≥325mg/day) at 48 hours, corticosteroid injection or hyaluronic acid injections at 3 and 6 months respectively. The use of antiulcer drugs, including histamine-2 receptor antagonists, proton pump inhibitors, misoprostol, and sucralfate was discontinued at least 24 hours before the baseline endoscopy.. Indirectness: No indirectness Comments: Valdecoxib is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=205) Intervention 3: Placebo. Placebo twice daily. Duration 12 weeks. Concurrent medication/care: People discontinued their normal medications at the following specified times before the baseline endoscopy: NSAIDs (including full dose aspirin at a dosage ≥325mg/day) at 48 hours, corticosteroid injection or hyaluronic acid injections at 3 and 6 months respectively. The use of antiulcer drugs, including histamine-2 receptor antagonists, proton pump inhibitors, misoprostol, and sucralfate was discontinued at least 24 hours before the baseline endoscopy.. Indirectness: No indirectness
Funding	Study funded by industry (This work was sponsored by Pharmacia Corporation and Pfizer, Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Patient's assessment of arthritis pain at 12 weeks; Group 1: mean -31.83 (SD 29.66); n=204, Group 2: mean -25.97 (SD 29.59); n=205; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports mean change scores and confidence intervals. Converted to SD. Reported naproxen: -31.83 (-35.90 to -27.76). Reported placebo: -25.97 (-30.02 to -21.92). Baseline naproxen: 72.36. Baseline placebo: 71.20.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, race, sex, disease

duration, history of GI bleeding and gastroduodenal ulcers, and baseline values for outcomes; Group 1 Number missing: 56, Reason: 56 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up; Group 2 Number missing: 74, Reason: 74 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastroduodenal ulcers at the final endoscopic evaluation at 12 weeks; Group 1: 18/183, Group 2: 8/178; Comments: Naproxen: Gastric ulcers = 16, duodenal ulcers = 2, of those the number of symptomatic ulcers was 7. Placebo: Gastric ulcers = 8, duodenal ulcers = 0, of those the number of symptomatic ulcers was 0. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, race, sex, disease duration, history of GI bleeding and gastroduodenal ulcers, and baseline values for outcomes; Group 1 Number missing: 22, Reason: 22 excluded from the safety analysis. 56 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up; Group 2 Number missing: 27, Reason: 27 excluded from the safety analysis. 74 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Abnormal hepatic function at 12 weeks; Group 1: 0/183, Group 2: 0/178 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, race, sex, disease duration, history of GI bleeding and gastroduodenal ulcers, and baseline values for outcomes; Group 1 Number missing: 22, Reason: 22 excluded from the safety analysis. 56 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up; Group 2 Number missing: 27, Reason: 27 excluded from the safety analysis. 74 in total withdrew due to treatment failure, preexisting protocol violations, noncompliance, or adverse signs and symptoms, or were lost to follow-up	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Kivitz 2004[103]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1042)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A diagnosis of osteoarthritis of the knee for longer than 6 months with a positive history of therapeutic response to NSAIDs and an American College of Rheumatology rating of I-III.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and nonpregnant females aged 40 and older with a diagnosis of osteoarthritis of the knee for longer than 6 months with a positive history of therapeutic response to NSAIDs and an American College of Rheumatology rating of I-III. After withdrawal of previous NSAID therapy, people were required to meet flare criteria at visit 2, which included minimum 40mm out of 100mm on pain walking assessment, increase of 15mm on pain walking assessment from visit 1 and worsening of IGADS by 1 point on a 5 point scale from visit 1.
Exclusion criteria	Concurrent medical/arthritic disease that could alter study outcome; a significant systemic disease that contraindicated NSAID therapy; people who used corticosteroids, misoprostol, sucralfate, histamine blockers, antacids, proton pump inhibitors, analgesics, warfarin, ticlopidine, high-dose aspirin, appetite suppressants, and other medications for chronic diseases for a predefined period before the study or if their use was required during the trial.
Recruitment/selection of patients	After withdrawal of previous NSAID therapy, people were required to meet flare criteria at visit 2, which included minimum 40mm out of 100mm on pain walking assessment, increase of 15mm on pain walking assessment from visit 1 and worsening of IGADS by 1 point on a 5 point scale from visit 1.
Age, gender and ethnicity	Age - Mean (SD): 63.2 (10.2). Gender (M:F): 330:712. Ethnicity: Caucasian = 916. The ethnicity of other groups were not defined.
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: American College of Rheumatology functional class I-III, median class II Duration of symptoms (mean [SD]): 6.1 (7.3) years
Indirectness of population	No indirectness
Interventions	(n=410) Intervention 1: NSAIDs - Other. Nabumetone 1000mg (as two 500mg tablets) once daily and matching placebo (for rofecoxib). Duration 6 weeks. Concurrent medication/care: Throughout the study, people were allowed to use paracetamol up to 2600mg a day as rescue medication for osteoarthritis pain, except during the first 6 days of therapy. People were also instructed not to use paracetamol for 24 hours before all efficacy evaluations.. Indirectness: No indirectness (n=423) Intervention 2: NSAIDs - Other. Rofecoxib 12.5mg once daily with matching placebo. Duration 6 weeks. Concurrent medication/care: Throughout the study, people were allowed to use paracetamol up to 2600mg a day as rescue medication for osteoarthritis pain, except during the first 6 days of therapy. People were also instructed not to use paracetamol for 24 hours before all efficacy evaluations.. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=208) Intervention 3: Placebo. Matching placebo twice a day. Duration 6 weeks. Concurrent medication/care: Throughout the study, people were allowed to use paracetamol up to 2600mg a day as rescue medication for osteoarthritis pain, except during the first 6 days of therapy. People were also instructed not to use paracetamol for 24 hours before all efficacy evaluations.. Indirectness: No indirectness
Funding	Study funded by industry (Research supported by Merck & Co., Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NABUMETONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Walking pain score at 6 weeks; MD; -11.4 (95%CI -15.5 to -7.3) (P-value: <0.001) Visual analogue scale 0-100 Top=High is poor outcome, Comments: Baseline nabumetone: 74.2 (15.3). Baseline placebo: 76.5 (14.6). N nabumetone = 410, N placebo = 208.;
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, duration of osteoarthritis, outcome baseline values, American College of Rheumatology functional class, NSAID-related gastrointestinal disease, prior NSAID use, prior paracetamol use, prior aspirin use; Group 1 Number missing: 85, Reason: Discontinued = 85 (20.7%). Lack of efficacy = 47, clinical adverse event = 25, withdrew consent = 4, protocol deviation = 4, lost to follow up = 1, other = 3, laboratory adverse event = 1; Group 2 Number missing: 67, Reason: Discontinued

= 67 (32.2%). Lack of efficacy = 49, clinical adverse event = 6, withdrew consent = 5, protocol deviation = 6, laboratory adverse event = 1

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal nuisance AEs (including acid reflux, dyspepsia, epigastric discomfort, heartburn, nausea, vomiting) at 6 weeks; Group 1: 21/410, Group 2: 14/208; Comments: Nabumetone: At least 1 AE = 21, acid reflux = 1, dyspepsia = 5, epigastric discomfort = 2, heartburn = 10, nausea = 5, vomiting = 1. Placebo: At least 1 AE = 14, acid reflux = 0, dyspepsia = 1, epigastric discomfort = 2, heartburn = 2, nausea = 9, vomiting = 1
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, duration of osteoarthritis, outcome baseline values, American College of Rheumatology functional class, NSAID-related gastrointestinal disease, prior NSAID use, prior paracetamol use, prior aspirin use; Group 1 Number missing: 85, Reason: Discontinued = 85 (20.7%). Lack of efficacy = 47, clinical adverse event = 25, withdrew consent = 4, protocol deviation = 4, lost to follow up = 1, other = 3, laboratory adverse event = 1; Group 2 Number missing: 67, Reason: Discontinued = 67 (32.2%). Lack of efficacy = 49, clinical adverse event = 6, withdrew consent = 5, protocol deviation = 6, laboratory adverse event = 1

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Lower extremity oedema at 6 weeks; Group 1: 7/410, Group 2: 2/208
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, duration of osteoarthritis, outcome baseline values, American College of Rheumatology functional class, NSAID-related gastrointestinal disease, prior NSAID use, prior paracetamol use, prior aspirin use; Group 1 Number missing: 85, Reason: Discontinued = 85 (20.7%). Lack of efficacy = 47, clinical adverse event = 25, withdrew consent = 4, protocol deviation = 4, lost to follow up = 1, other = 3, laboratory adverse event = 1; Group 2 Number missing: 67, Reason: Discontinued = 67 (32.2%). Lack of efficacy = 49, clinical adverse event = 6, withdrew consent = 5, protocol deviation = 6, laboratory adverse event = 1

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 6 weeks; Group 1: 36/410, Group 2: 28/208
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, ethnicity, age, duration of osteoarthritis, outcome baseline values, American College of Rheumatology functional class, NSAID-related gastrointestinal disease, prior NSAID use, prior paracetamol use, prior aspirin use; Group 1 Number missing: 85, Reason: Discontinued = 85 (20.7%). Lack of efficacy = 47, clinical adverse event = 25, withdrew consent = 4, protocol deviation = 4, lost to follow up = 1, other = 3, laboratory adverse event = 1; Group 2 Number missing: 67, Reason: Discontinued = 67 (32.2%). Lack of efficacy = 49, clinical adverse event = 6, withdrew consent = 5, protocol deviation = 6, laboratory adverse event = 1

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Kneer 2013[106]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=867)
Countries and setting	Conducted in Croatia, Germany, Poland, Serbia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of osteoarthritis in at least one knee for a minimum of 6 months, meeting at least two of the following: morning stiffness lasting <30 minutes, crepitus on motion, or age at least 40 years. People had to meet the American College of Rheumatology clinical criteria for osteoarthritis and have an American College of Rheumatology functional class rating of I-III. The radiographic criteria for the index knee were a Kellgren-Lawrence score of grade 2-3 with radiographs taken within 6 months before baseline
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 18-75 years with a clinical diagnosis of osteoarthritis in at least one knee for a minimum of 6 months, meeting at least two of the following: morning stiffness lasting <30 minutes, crepitus on motion, or age at least 40 years. People had to meet the American College of Rheumatology clinical criteria for osteoarthritis and have an American College of Rheumatology functional class rating of I-III. The radiographic criteria for the index knee were a Kellgren-Lawrence score of grade 2-3 with radiographs taken within 6 months before baseline. In addition, people had to be able to walk at least 100 feet without an assistive device. The people had to have used a daily dose of an oral or rectal NSAID on at least 3 days per week during the 3 months before screening or on at least 25 of the 30 days before screening and had to be dissatisfied with their current NSAID treatment.
Exclusion criteria	The presence of any causes of secondary osteoarthritis; diseases of the spine or lower extremity joints potentially affecting the assessment of the index knee; severe coexisting diseases, such as peptic ulcers; severe renal, cardiovascular, or neurologic diseases; use of intra-articular medications or arthroscopy within the preceding 3 months; and skin lesions or dermatologic diseases in the treatment area.
Recruitment/selection of patients	People were required to stop treatment with their current NSAID and to return for a baseline visit at the end of a washout period. The length of the washout period was determined by the half-life of the patient's NSAID and lasted for five half-lives plus 2 additional days. At the baseline (randomisation) visit, the osteoarthritis flare criteria for

	the index knee were evaluated: people with a WOMAC VAS pain subscale score of >40mm and an increase of at least 15mm as compared with the value at the screening visit before stopping NSAID treatment were eligible for randomisation.
Age, gender and ethnicity	Age - Mean (SD): 61.7 (9.3). Gender (M:F): 235:593. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range (19-78)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not explicitly stated. At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=667) Intervention 1: NSAID gels (topical - local) - Other. Ketoprofen gel including either 25, 50 or 100mg of ketoprofen in a Transfersome gel. Applied twice daily for 12 weeks. Dosing was controlled by a dispensing device with each dose being equivalent to one stroke of the respective gel. The gel was spread gently and homogenously over the knee(s) and around the knee, including the popliteal fossa, but sparing the patella. The upper boundary extended around the leg, from approximately 5cm above the superior edge of the patella. The lower boundary extended around the leg from the inferior edge of the tibial tuberosity (approximately 5cm below the inferior edge of the patella). In order to prevent any confounding effects from excessive massage, the people were instructed that rubbing, kneading and massaging had to be avoided. The gel had to dry for at least 15 minutes before putting on clothes.. Duration 12 weeks. Concurrent medication/care: Paracetamol, up to a maximum daily dose of 2g/day for up to 5 days during any 7-day period, were permitted for breakthrough pain or non-OA pain. Rescue medication use was not allowed within 48 hours before the study visits.. Indirectness: No indirectness Comments: The groups with different doses of ketoprofen were pooled together for analysis as agreed in the protocol (n=199) Intervention 2: Placebo. Transfersome gel only (no ketoprofen). Applied twice daily for 12 weeks. Dosing was controlled by a dispensing device with each dose being equivalent to one stroke of the respective gel. The gel was spread gently and homogenously over the knee(s) and around the knee, including the popliteal fossa, but sparing the patella. The upper boundary extended around the leg, from approximately 5cm above the superior edge of the patella. The lower boundary extended around the leg from the inferior edge of the tibial tuberosity (approximately 5cm below the inferior edge of the patella). In order to prevent any confounding effects from excessive massage, the people were instructed that rubbing, kneading and massaging had to be

	avoided. The gel had to dry for at least 15 minutes before putting on clothes.. Duration 12 weeks. Concurrent medication/care: Paracetamol, up to a maximum daily dose of 2g/day for up to 5 days during any 7-day period, were permitted for breakthrough pain or non-OA pain. Rescue medication use was not allowed within 48 hours before the study visits.. Indirectness: No indirectness
Funding	Principal author funded by industry (Werner Kneer and Egbert J Seidel received investigator grants from the sponsor, IDEA AG. Matthias Rother and Stefan Mazgareanu were employees of IDEA AG during the conduct of the study. Matthias Rother is a paid consultant of Pro Bono Bio Entrepreneur Ltd)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean 28.73 (SD 21.16); n=638, Group 2: mean 32.57 (SD 32.33); n=190; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reported ketoprofen 100mg: 28.39 (21.00). Reported ketoprofen 50mg: 27.92 (21.28). Reported ketoprofen 25mg: 29.88 (21.16). Reported placebo: 32.57 (32.33). Baseline ketoprofen 100mg: 65.67 (13.38). Baseline ketoprofen 50mg: 65.35 (14.35). Baseline ketoprofen 25mg: 64.10 (13.90). Baseline placebo: 65.6 (13.2).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 12 weeks; Group 1: mean 30.25 (SD 20.78); n=638, Group 2: mean 33.16 (SD 21.75); n=190; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Reported ketoprofen 100mg: 30.56 (21.44). Reported ketoprofen 50mg: 29.07 (21.20). Reported ketoprofen 25mg: 32.12 (19.62). Reported placebo: 33.16 (21.75). Baseline ketoprofen 100mg: 52.85 (17.38). Baseline ketoprofen 50mg: 53.05 (18.25). Baseline ketoprofen 25mg: 51.21 (18.49). Baseline placebo: 52.54 (16.79).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders at 12 weeks; Group 1: 22/667, Group 2: 9/199; Comments: Ketoprofen 100mg: 10, Ketoprofen 50mg: 6, Ketoprofen 25mg: 6, Placebo: 9

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiac disorders at 12 weeks; Group 1: 8/667, Group 2: 1/199; Comments: Ketoprofen 100mg: 3, Ketoprofen 50mg: 2, Ketoprofen 25mg: 3, Placebo: 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Renal and urinary disorders at 12 weeks; Group 1: 13/667, Group 2: 1/199; Comments: Ketoprofen 100mg: 4, Ketoprofen 50mg: 4, Ketoprofen 25mg: 5, Placebo: 1 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system disorders at 12 weeks; Group 1: 50/667, Group 2: 23/199; Comments: Ketoprofen 100mg: 19, Ketoprofen 50mg: 17, Ketoprofen 25mg: 14, Placebo: 23 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, knee affected, and WOMAC baseline values; Group 1 Number missing: 112, Reason: Ketoprofen 100mg: 16 adverse events, 6 insufficient effect, 13 other. Ketoprofen 50mg: 15 adverse events, 14 insufficient effect, 9 other. Ketoprofen 25mg: 9 adverse events, 15 insufficient effect, 15 other.; Group 2 Number missing: 36, Reason: 8 adverse events, 12 insufficient effect, 16 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Kosuwon 2010[108]
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	1 (n=99)
Countries and setting	Conducted in Thailand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 9 weeks (two 4 week treatment phases, and a 1 week washout period)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A diagnosis of osteoarthritis of the knee for at least 6 months prior to screening who have met the American College of Rheumatology clinical criteria for the classification of idiopathic (primary) osteoarthritis. Documented radiographic evidence of osteoarthritis of the knee of grade 2 or 3 on the Kellgren and Lawrence radiographic grade.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People capable of giving written informed consent; ambulatory males or non-pregnant females between 40 and 80 years of age; off pain medication or nutritional supplements for symptomatic relief of knee OA at least 15 days before screening; pain in the knee ≤80mm on a 100mm VAS; have a diagnosis of osteoarthritis of the knee; a baseline minimum joint space width in the medial and lateral compartments of the index knee of ≥1.5 and ≥2.5mm respectively, measured from radiographs in the MTP view
Exclusion criteria	History of hypersensitivity to capsaicin; skin lesion at the index knee; a history of lower extremity surgery within 6 months prior to screening; significant prior injury to the index knee within 12 months prior to screening; disease of the spine or other lower extremity joints of sufficient degree to affect the index knee; treatment with other drugs potentially affecting bone or cartilage metabolism, such as: chronic systematic corticosteroid, hyaluronan injection into the index knee within the previous 6 months and diacerin treatment within the last 12 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 61 (44-82). Gender (M:F): 0:99. Ethnicity: Not reported
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Kellgren Lawrence grade 2-3. Duration of symptoms: Not stated explicitly. At least 6 months..
Indirectness of population	No indirectness
Interventions	(n=99) Intervention 1: Capsaicin cream (topical - local) - Capsaicin cream (topical). Capsicum tincture 45.50g (equivalent to capsaicin 0.0125%) per 100g of Capsika gel® - 2 inches applied around the index knee and rubbed in until dry three times a day. Duration 4 weeks. Concurrent medication/care: People were permitted to take paracetamol for pain (500mg three times a day or every 4-6 hours) but not any other topical analgesics, NSAIDs or COX-2 inhibitors. They worked as usual and the authors did not provide any knee brace or physical therapy.. Indirectness: No indirectness (n=99) Intervention 2: Placebo. Placebo gel, applied in the same way as capsaicin. Duration 4 weeks. Concurrent medication/care: People were permitted to take paracetamol for pain (500mg three times a day or every 4-6 hours) but not any other topical analgesics, NSAIDs or COX-2 inhibitors. They worked as usual and the authors did not provide any knee brace or physical therapy.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CAPSAICIN CREAM (TOPICAL) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 4 weeks; Group 1: mean 4.66 (SD 4.14); n=99, Group 2: mean 1.24 (SD 3.55); n=99; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline: 11.09 (3.06)
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 4 weeks; Group 1: mean 14.54 (SD 13.62); n=99, Group 2: mean 5.56 (SD 10.79); n=99; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline: 38.92 (9.89).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised,

of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out. Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Other adverse events at 4 weeks; Group 1: 0/99, Group 2: 0/99; Comments: "A burning sensation was the only adverse event reported in our study" Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Other adverse events at 4 weeks; Group 1: 0/99, Group 2: 0/99; Comments: "A burning sensation was the only adverse event reported in our study" Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out. Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Other adverse events at 4 weeks; Group 1: 0/99, Group 2: 0/99; Comments: "A burning sensation was the only adverse event reported in our study" Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out. Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Other adverse events at 4 weeks; Group 1: 0/99, Group 2: 0/99; Comments: "A burning sensation was the only adverse event reported in our study" Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall (cross over study) values for sex, age, affected knee, weight and height, joint space and baseline values of outcomes; Group 1 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.; Group 2 Number missing: 0, Reason: Originally 100 people were randomised, of which 1 dropped out.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Kwoh 2014[109]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=201)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Mild-to-moderate chronic, frequent knee pain typical of knee osteoarthritis (WOMAC score of at least 25 but less than 100)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People between the ages of 35 years and 65 years with symptoms of mild-to-moderate chronic, frequent knee pain typical of osteoarthritis
Exclusion criteria	People with Kellgren-Lawrence grade of 4 on radiographs of both knees; inflammatory arthritis; renal disease; liver disease; diabetes mellitus; cancer; plans for elective surgery in the subsequent 12 months; an inability to undergo MRI of the knee; inability to walk without a cane or other assistive devices; people who were taking bisphosphonates or who had taken glucosamine or other dietary supplements for knee pain within the 6 months prior; people unwilling to avoid treatment of knee pain with the use of NSAIDs or any pain relievers other than paracetamol for 24 weeks
Recruitment/selection of patients	People were recruited from the community through mass mailings and physician offices, and used the University of Pittsburgh Arthritis Registry with a telephone screening interview and an in-person screening visit at the Arthritis Research Clinic at the University of Pittsburgh
Age, gender and ethnicity	Age - Mean (SD): 52.23 (6.40). Gender (M:F): 103:98. Ethnicity: White = 184, others not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Mixed (Includes people with Kellgren Lawrence grade 0 changes). 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Seveirty: Kellgren Lawrence grade 0-4 Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=98) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 1500mg in a 16-ounce bottle of diet

	lemonade. Non-nutritive sweeteners and flavourings (e.g. sucra-lose and acesulfame K) were used to minimize the provision of additional calories and to ensure similar taste.. Duration 24 weeks. Concurrent medication/care: Paracetamol was the only analgesic allowed during the study. Indirectness: No indirectness (n=103) Intervention 2: Placebo. A 16-ounce bottle of diet lemonade. Non-nutritive sweeteners and flavourings (e.g. sucra-lose and acesulfame K) were used to minimize the provision of additional calories and to ensure similar taste.. Duration 24 weeks. Concurrent medication/care: Paracetamol was the only analgesic allowed during the study. Indirectness: No indirectness
Funding	Study funded by industry (Funding for this study was provided by the Beverage Institute for Health & Wellness)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain at 12 weeks; Group 1: mean -20.0714 (SD 17.31); n=98, Group 2: mean -20.0893 (SD 21.33); n=103; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 45.20 (14.03). Baseline placebo: 47.72 (16.65).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, WOMAC baseline score, use of analgesics, radiographic features and MRI features; Group 1 Number missing: 5, Reason: 1 disliked taste, 3 side effects, 1 knee procedure; Group 2 Number missing: 17, Reason: 5 lost to follow up, 2 disliked taste, 2 illness, 4 side effects, 4 others
- Actual outcome for Knee: WOMAC pain at 24 weeks; Group 1: mean -17.402 (SD 20.99); n=98, Group 2: mean -20.8893 (SD 21.43); n=103; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 45.20 (14.03). Baseline placebo: 47.72 (16.65).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, WOMAC baseline score, use of analgesics, radiographic features and MRI features; Group 1 Number missing: 5, Reason: 1 disliked taste, 3 side effects, 1 knee procedure; Group 2 Number missing: 17, Reason: 5 lost to follow up, 2 disliked taste, 2 illness, 4 side effects, 4 others

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC difficulty at 12 weeks; Group 1: mean -18.0873 (SD 17.85); n=98, Group 2: mean -18.718 (SD 22.3); n=103; WOMAC difficulty 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 42.76 (16.89). Baseline placebo: 47.34 (19.60).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, WOMAC baseline score, use of analgesics, radiographic features and MRI features; Group 1 Number missing: 5, Reason: 1 disliked taste, 3 side effects, 1 knee procedure; Group 2 Number missing: 17, Reason: 5 lost to follow up, 2 disliked taste, 2 illness, 4 side effects, 4 others
- Actual outcome for Knee: WOMAC difficulty at 24 weeks; Group 1: mean -15.4138 (SD 21.34); n=98, Group 2: mean -19.404 (SD 21.21); n=103; WOMAC

function subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 42.76 (16.89). Baseline placebo: 47.34 (19.60). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, WOMAC baseline score, use of analgesics, radiographic features and MRI features; Group 1 Number missing: 5, Reason: 1 disliked taste, 3 side effects, 1 knee procedure; Group 2 Number missing: 17, Reason: 5 lost to follow up, 2 disliked taste, 2 illness, 4 side effects, 4 others	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Laine 1999[110]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=741)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 weeks
Method of assessment of guideline condition	Method of assessment /diagnosis not stated: States people with osteoarthritis
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis over 50 years old requiring NSAID treatment for at least 6 months. People with baseline gastroduodenal erosions at baseline were eligible.
Exclusion criteria	Active duodenal, gastric, and oesophageal ulcers; pyloric obstruction; erosive oesophagitis at baseline endoscopy; past upper GI surgery; inflammatory bowel disease; serum creatinine level of >2.0mg/dL; an estimated creatinine clearance of ≤30mL/min; faecal occult blood; unstable medical disease; a history of malignancy in the prior 5 years; cerebrovascular events in the prior 2 years; a bleeding diathesis; a requirement for anticoagulant therapy, corticosteroids, ticlopidine or aspirin.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 61 (47-87). Gender (M:F): 241:500. Ethnicity: White = 331, Black = 41, Hispanic = 20, Other = 8
Further population details	1. Age: Mixed 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Unclear).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=184) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 800mg three times daily. Duration 24 weeks. Concurrent medication/care: People were permitted to take supplied paracetamol (up to 2600mg daily), non-NSAID pain medications and supplied antacid as needed. They were not permitted to take nonstudy NSAIDs, aspirin, corticosteroids, anticoagulants, ticlopidine, histamine 2 receptor antagonists, sucralfate, prostaglandin analogues, proton pump inhibitors or antacids other than Gelusil.. Indirectness: No indirectness

	(n=381) Intervention 2: NSAIDs - Other. Rofecoxib 25mg or 50mg once daily. Duration 24 weeks. Concurrent medication/care: People were permitted to take supplied paracetamol (up to 2600mg daily), non-NSAID pain medications and supplied antacid as needed. They were not permitted to take nonstudy NSAIDs, aspirin, corticosteroids, anticoagulants, ticlopidine, histamine 2 receptor antagonists, sucralfate, prostaglandin analogues, proton pump inhibitors or antacids other than Gelusil.. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK so was not included in the analysis, as agreed in the protocol. It is reported here for completeness. (n=177) Intervention 3: Placebo. Matching placebo. Duration 24 weeks. Concurrent medication/care: People were permitted to take supplied paracetamol (up to 2600mg daily), non-NSAID pain medications and supplied antacid as needed. They were not permitted to take nonstudy NSAIDs, aspirin, corticosteroids, anticoagulants, ticlopidine, histamine 2 receptor antagonists, sucralfate, prostaglandin analogues, proton pump inhibitors or antacids other than Gelusil.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by a grant from Merck & Co., Inc.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN versus PLACEBO Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Gastroduodenal ulcer incidence and upper GI bleeding episodes at 12 weeks; Group 1: 48/167, Group 2: 16/158; Comments: Reported ibuprofen: 27.7% (95% CI 20.4-35.0) and 2 upper GI bleeding episodes. Reported placebo: 9.9% (95% CI 4.1-15.7). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, history of GI events, H. pylori status, Patient's global assessment of disease mean score, presence of baseline gastroduodenal erosions, tobacco and alcohol use; Group 1 Number missing: 17, Reason: 16 discontinued prior to week 6 and had no repeat endoscopy. 1 person was allocated serially by 2 different study centers and had their second enrollment period excluded from analysis.; Group 2 Number missing: 19, Reason: 19 discontinued prior to week 6 and had no repeat endoscopy	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Langford 2006[114]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=399)
Countries and setting	Conducted in Canada, Czech Republic, Hungary, Poland, Slovakia, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (1 week pretreatment run in phase)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Meeting the American College of Rheumatology diagnostic criteria for hip or knee osteoarthritis and requiring joint replacement surgery, with radiographic evidence of disease in the affected joints
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People (at least 40 years of age) meeting the American College of Rheumatology diagnostic criteria for hip or knee osteoarthritis and requiring joint replacement surgery, with radiographic evidence of disease in the affected joint(s). People could be included in the study if they were awaiting surgery, if they had refused surgery, or if they were unable to undergo surgery for medical reasons. All people had experienced joint pain for more than 3 months, and for at least 20 days each month. All people had to have moderate or severe pain that was not adequately controlled with weak opioids, with or without paracetamol. Pain was assessed each morning and evening using a 100mm visual analogue scale (higher scores indicate greater pain) in response to questions. To be eligible for the study, people had to have mean daily visual analogue scale pain score of at least 50 at the start and end of the 7 days of the pretreatment run-in phase prior to initiation of treatment and a mean VAS pain score of at least 50 for the entire 7 days of the pretreatment phase.
Exclusion criteria	Received any strong opioid in the 4 weeks before the study or had recently started a new therapy (e.g. physiotherapy or acupuncture) for their pain. Those people deemed unsuitable for treatment with a strong opioid (e.g. because of suspected alcohol or drug abuse, or because they were considered at risk for respiratory depression).
Recruitment/selection of patients	People had to have a washout of medications and an increase in pain after washout
Age, gender and ethnicity	Age - Mean (range): 66 (40-90). Gender (M:F): 234:165. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).

Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 3 months.
Indirectness of population	No indirectness
Interventions	(n=216) Intervention 1: Opioids (topical - systemic) - Opioids (topical systemic). Transdermal fentanyl. 1 week run in with a dosage of 25 micrograms/hour replaced every 72 hours. Increased to higher doses (to a maximum of 100 micrograms/hour) every 3 days.. Duration 6 weeks. Concurrent medication/care: People were asked to continue to receive stable doses of antiinflammatory agents (steroids or NSAIDs, including COX-2 inhibitors) that were prescribed before the study, but all weak opioids were stopped. People could also take up to 4 grams of paracetamol per day (but not combination preparations of paracetamol and weak opioids). People were encouraged to take metoclopramide (supplied as 10mg tablets) immediately if they experienced any nausea or vomiting. They were also encouraged to take a laxative if they had constipation.. Indirectness: No indirectness (n=200) Intervention 2: Placebo. Matching placebo patches. Duration 6 weeks. Concurrent medication/care: People were asked to continue to receive stable doses of antiinflammatory agents (steroids or NSAIDs, including COX-2 inhibitors) that were prescribed before the study, but all weak opioids were stopped. People could also take up to 4 grams of paracetamol per day (but not combination preparations of paracetamol and weak opioids). People were encouraged to take metoclopramide (supplied as 10mg tablets) immediately if they experienced any nausea or vomiting. They were also encouraged to take a laxative if they had constipation.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by funds from Janssen-Cilag)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OPIOIDS (TOPICAL SYSTEMIC) versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical functioning at 6 weeks; Group 1: mean 4.8 (SD 17.1); n=202, Group 2: mean 2.9 (SD 18.3); n=197; SF-36 physical functioning subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 4.8 (1.2). Reported placebo: 2.9 (1.3). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow

up, 7 other
- Actual outcome for Other: SF-36 role-physical at 6 weeks; Group 1: mean 5.3 (SD 39.8); n=202, Group 2: mean 7.8 (SD 33.7); n=197; SF-36 role-physical subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 5.3 (2.8). Reported placebo: 7.8 (2.4). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 pain index at 6 weeks; Group 1: mean 11.4 (SD 19.9); n=202, Group 2: mean 7.1 (SD 19.6); n=197; SF-36 pain index subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 11.4 (1.4). Reported placebo: 7.1 (1.4). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 general health at 6 weeks; Group 1: mean 2.4 (SD 17.1); n=202, Group 2: mean 3.4 (SD 15.4); n=197; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 2.4 (1.2). Reported placebo: 3.4 (1.1). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 vitality at 6 weeks; Group 1: mean 1.9 (SD 21.3); n=202, Group 2: mean 3.1 (SD 19.7); n=197; SF-36 vitality 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 1.9 (1.5). Reported placebo: 3.1 (1.4). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 social functioning at 6 weeks; Group 1: mean 3.2 (SD 34.1); n=202, Group 2: mean 6.3 (SD 26.7); n=197; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: 3.2 (2.4). Reported placebo: 6.3 (1.9). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes.

Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 role-emotional at 6 weeks; Group 1: mean -2.4 (SD 52.6); n=202, Group 2: mean 6 (SD 42.1); n=197; SF-36 role-emotional subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: -2.4 (3.7). Reported placebo: 6.0 (3.0). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other
- Actual outcome for Other: SF-36 mental health at 6 weeks; Group 1: mean -0.4 (SD 19.9); n=202, Group 2: mean 0.7 (SD 16.8); n=197; SF-36 mental health subscale 0-100 Top=High is good outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: -0.4 (1.4). Reported placebo: 0.7 (1.2). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 6 weeks; Group 1: mean -1.5 (SD 1.4); n=202, Group 2: mean -0.8 (SD 1.4); n=197; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: -1.5 (0.1). Reported placebo: -0.8 (0.1). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow up, 7 other

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 6 weeks; Group 1: mean -1.1 (SD 1.4); n=202, Group 2: mean -0.7 (SD 1.4); n=197; WOMAC physical function subscale Range unclear Top=High is poor outcome; Comments: Reports change score and standard error. Converted into standard deviation. Reported fentanyl: -1.1 (0.1). Reported placebo: -0.7 (0.1). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline values for outcomes. Reports age, weight, gender and baseline pain score (VAS).; Group 1 Number missing: 96, Reason: 54 adverse events, 15 insufficient efficacy, 17 withdrew consent, 1 lost to follow-up, 9 other; Group 2 Number missing: 104, Reason: 20 adverse events, 64 insufficient efficacy, 13 withdrew consent, 0 lost to follow

up, 7 other	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Leatham 1983[115]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=36)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 9 weeks (3 weeks for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip and/or knee as judged by the presence of at least 2 of the following 3 radiological criteria: narrowing of joint space, presence of osteophytes, osteosclerosis at the joint margin
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the hip and/or knee with pain at rest, pain on standing or pain on walking. All people had a normal ESR, negative latex fixation test and a normal blood uric acid.
Exclusion criteria	People with radiological evidence of pyrophosphate deposition in articular cartilage; pregnancy or females of child bearing age; chronic hepatic or chronic renal insufficiency; presence of congestive cardiac failure; history of previous peptic ulceration
Recruitment/selection of patients	People had a two week washout period on paracetamol alone before allocation to the treatment arms
Age, gender and ethnicity	Age - Other: Not reported. Gender (M:F): Not reported. Ethnicity: Not stated
Further population details	1. Age: Not stated / Unclear 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip and/or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=36) Intervention 1: NSAIDs - Naproxen. Naproxen 250mg twice daily with two placebo tablets twice daily. Duration 3 weeks. Concurrent medication/care: No other analgesics or anti-rheumatic drugs were allowed except paracetamol alone as 'back up' therapy. Indirectness: No indirectness (n=36) Intervention 2: NSAIDs - Other. Antrafenine 150mg four times a day. Duration

	3 weeks. Concurrent medication/care: No other analgesics or anti-rheumatic drugs were allowed except paracetamol alone as 'back up' therapy. Indirectness: No indirectness Comments: Antrafenine is not licensed for use in the UK and so was not included in this analysis as agreed in the protocol. This was reported here for completeness. (n=36) Intervention 3: Placebo. Placebo four times a day. Duration 3 weeks. Concurrent medication/care: No other analgesics or anti-rheumatic drugs were allowed except paracetamol alone as 'back up' therapy. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Other: Pain on walking (visual analogue scale) at 3 weeks; Group 1: mean 2.9 (SD 1); n=28, Group 2: mean 3.7 (SD 1.1); n=28; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reports mean and standard error. Calculated SD from this. Reported naproxen: 2.9 (0.18). Reported placebo: 3.7 (0.20). Overall baseline: 3.4 (0.21) Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report any baseline characteristics apart from baseline value for outcome; Group 1 Number missing: 8, Reason: 2 withdrew due to lack of efficacy on placebo, 3 withdrew because they failed to attend the clinic, 3 failed to complete the trial due to concomitant illness; Group 2 Number missing: 8, Reason: 2 withdrew due to lack of efficacy on placebo, 3 withdrew because they failed to attend the clinic, 3 failed to complete the trial due to concomitant illness	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Lee 2017[116]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=362)
Countries and setting	Conducted in South Korea; Setting: Tertiary care follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (followed by an 18 week single arm open label extension)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of knee or hip osteoarthritis according to clinical and imaging criteria specified by the American College of Rheumatology guidelines
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Males or females aged at least 20 years with a clinical diagnosis of knee or hip osteoarthritis according to clinical and imaging criteria specified by the American College of Rheumatology guidelines. People must have had chronic pain for at least 3 months from osteoarthritis and an American College of Rheumatology global functional status of I-III (excluding IV). Other inclusion criteria were as follows: screening and baseline mean WOMAc pain score in the index joint between 4 and 8; blood chemistry within twice the normal range; urinalysis within normal limits, macroscopic evaluation showing no blood; agreement to use double barrier contraception during the study period and for 3 months afterwards; able to read, understand and follow the study documents; willing to limit alcohol intake for at less than or equal to 2 drinks per day. People with stable hypertension for at least 3 months prior to screening, treated with antihypertensive medication.
Exclusion criteria	Use of any analgesics except the study medications or paracetamol during the study; use of anticoagulants within 2 weeks of screening; use of corticosteroids, herbal medicines, traditional Korean medicines, neutraceuticals, glucosamine and/or chondroitin sulfate; requirement for knee or hip arthroplasty within 2 months of screening or anticipating any need for a surgical procedure on the index joint during the study; hypersensitivity or allergy to NSAIDs; history of nasal polyps, bronchospasm, urticaria, or anaphylactic shock; history of New York Heart Association stage II-IV congestive heart failure, ischaemic heart disease, uncontrolled hypertension, peripheral arterial disease and/or cerebrovascular disease; pregnancy, breast-feeding, or expecting conception within the duration of the study; active ulcer, GI bleeding, ulcerative colitis, or severe renal, hepatic or anticoagulant disorder within

	6 months prior to randomisation; ongoing chronic symptoms or psychiatric disorders preventing compliance with study procedures, except for subjects who were physically healthy and had been receiving the specified allowed drugs for at least 3 months; use of corticosteroids, intra-articular steroids or hyaluronic acid injection within 1 month of screening; chemotherapy within 5 years; a Chinese traditional arthritis treatment within 1 week of screening.
Recruitment/selection of patients	Conducted at 14 tertiary care centers in Korea. People in the study entered a 14 day washout period where they were required to discontinue existing treatment with NSAIDs and/or other analgesic medication.
Age, gender and ethnicity	Age - Mean (SD): 62.4 (7.8). Gender (M:F): 53:309. Ethnicity: All participants were Asian
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: High comorbidity score (Hypertension in 119, hyperlipidaemia in 70, osteoporosis in 38, diabetes mellitus in 36, meniscus injury in 35, gastritis in 32). 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American College of Rheumatology global functional status of I-III Duration of symptoms: Not stated explicitly. At least 3 months..
Indirectness of population	No indirectness
Interventions	(n=145) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once daily. Duration 6 weeks. Concurrent medication/care: Rescue medication (paracetamol 650mg per day) was allowed during the washout and follow up period. Rescue medication was not allowed during the treatment period or within 24 hours before a clinic visit.. Indirectness: No indirectness (n=146) Intervention 2: NSAIDs - Other. Polmacoxib 2mg once a day. Duration 6 weeks. Concurrent medication/care: Rescue medication (paracetamol 650mg per day) was allowed during the washout and follow up period. Rescue medication was not allowed during the treatment period or within 24 hours before a clinic visit.. Indirectness: No indirectness Comments: Polmacoxib is not licensed for use in the UK and so was not included in the final analysis as agreed in the protocol. It is reported here for completeness. (n=71) Intervention 3: Placebo. Placebo once a day. Duration 6 weeks. Concurrent medication/care: Rescue medication (paracetamol 650mg per day) was allowed during the washout and follow up period. Rescue medication was not allowed during the treatment period or within 24 hours before a clinic visit.. Indirectness: No

	indirectness
Funding	Principal author funded by industry (Sangsook Cho was an employee of the Clinical Research Department of CG Phaemaceuticals Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 6 weeks; Group 1: mean -5.7 (SD 8.4); n=145, Group 2: mean -2.6 (SD 7.6); n=71; WOMAC pain subscale 0-50 Top=High is poor outcome; Comments: Reports standard error, converted to SD. Reported SE celecoxib: 0.7. Reported SE placebo: 0.9. Baseline celecoxib: 27.7 (5.08). Baseline placebo: 26.8 (4.58).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, smoking status, drinking status, concomitant medication use, medical condition or procedures (comorbidities), and WOMAC baseline values; Group 1 Number missing: 13, Reason: 4 protocol violations, 2 subject withdrawal, 5 adverse events, 2 lack of efficacy/other; Group 2 Number missing: 5, Reason: 4 subject withdrawal, 1 adverse event

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 6 weeks; Group 1: mean -14.9 (SD 26.3); n=145, Group 2: mean -7.9 (SD 23.9); n=71; WOMAC physical function subscale 0-170 Top=High is poor outcome; Comments: Reports standard errors, converted to SD. Reported SE celecoxib: 2.18. Reported SE placebo: 2.83. Baseline celecoxib: 92.8 (19.26). Baseline placebo: 92.0 (19.87).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, smoking status, drinking status, concomitant medication use, medical condition or procedures (comorbidities), and WOMAC baseline values; Group 1 Number missing: 13, Reason: 4 protocol violations, 2 subject withdrawal, 5 adverse events, 2 lack of efficacy/other; Group 2 Number missing: 5, Reason: 4 subject withdrawal, 1 adverse event

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	LEGS trial: Fransen 2015[73]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=605)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee pain for more than 6 months that is present on most days of the last month (or taking analgesics for the pain) with radiographic evidence using stand criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Knee pain for more than 6 months; knee pain on most days of the past month (or taking analgesics); worst pain experienced in the left and right knee over the past week rated as 4 or more on a scale of 0-10. People with diabetes or taking warfarin were provided with written advice encouraging increased vigilance with glucose or INR monitoring.
Exclusion criteria	Rheumatoid arthritis, bilateral knee replacements or unstable diabetes; allergy to shellfish, lower limb surgery in the past 6 months; planning to have a knee replacement in the next year; intra-articular injection for knee pain in the past 3 months
Recruitment/selection of patients	People were screened by telephone. Glucosamine was quality checked.
Age, gender and ethnicity	Age - Mean (SD): 60.5 (8.1). Gender (M:F): 266:339. Ethnicity: Not stated
Further population details	1. Age: <75 years (Based on SD). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Low comorbidity score (Mean (SD): 3.3 (3.0)). 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: 48-61% had less than Kellgren-Lawrence grade 2 osteoarthritis Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=152) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Two capsules of glucosamine sulfate potassium chloride once a day (753mg x 2) with two placebo capsules. Duration 2 years. Concurrent medication/care: No additional information. Indirectness: No indirectness

	(n=302) Intervention 2: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Either two capsules of glucosamine and two capsules of chondroitin, or two capsules of chondroitin (including 400mg x 2) and two placebo capsules once a day. Duration 2 years. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: Both the glucosamine and chondroitin combination and chondroitin only interventions were not included in the protocol and so are not included in the analysis. They are reported here for completeness. (n=151) Intervention 3: Placebo. Four placebo capsules once a day. Duration 2 years. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded through the National Health and Medical Research Council of Australia, the Department of Health and Aging, and by some supplementary funding (<15%) from Sanofi-Aventis Consumer Healthcare Pty Ltd, Australia. The funders had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: SF-12 physical component summary at 2 years; Group 1: mean 43.9 (SD 9.4); n=152, Group 2: mean 44.2 (SD 9.7); n=151; SF-12 physical component summary 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 40.9 (9.6). Baseline placebo: 42.1 (9.6).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, presence of Heberden's nodes, knee surgery history, comorbidity, diabetes, warfarin, hip pain, back pain, glucosamine use, radiographic severe, study knee JSW, previous analgesic use, and baseline values of outcomes; Group 1 Number missing: 27, Reason: 8 adverse events, 19 other; Group 2 Number missing: 34, Reason: 8 adverse events, 26 other
- Actual outcome for Knee: SF-12 mental component summary at 2 years; Group 1: mean 53.1 (SD 10.3); n=152, Group 2: mean 51.6 (SD 10); n=151; SF-12 mental component summary 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 53.1 (10.3). Baseline placebo: 51.6 (10.7).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, presence of Heberden's nodes, knee surgery history, comorbidity, diabetes, warfarin, hip pain, back pain, glucosamine use, radiographic severe, study knee JSW, previous analgesic use, and baseline values of outcomes; Group 1 Number missing: 27, Reason: 8 adverse events, 19 other; Group 2 Number missing: 34, Reason: 8 adverse events, 26 other

Protocol outcome 2: Pain reduction at ≤3- or >3- months

- Actual outcome for Knee: WOMAC pain subscale at 2 years; Group 1: mean 4.5 (SD 3.7); n=152, Group 2: mean 4.6 (SD 3.5); n=151; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline glucosamine: 6.5 (3.4). Baseline placebo: 6.7 (3.4). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, presence of Heberden's nodes, knee surgery history, comorbidity, diabetes, warfarin, hip pain, back pain, glucosamine use, radiographic severe, study knee JSW, previous analgesic use, and baseline values of outcomes; Group 1 Number missing: 27, Reason: 8 adverse events, 19 other; Group 2 Number missing: 34, Reason: 8 adverse events, 26 other Protocol outcome 3: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC function subscale at 2 years; Group 1: mean 17.8 (SD 13.5); n=152, Group 2: mean 17.8 (SD 12.9); n=151; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline glucosamine: 21.7 (11.8). Baseline placebo: 21.7 (12.8). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, presence of Heberden's nodes, knee surgery history, comorbidity, diabetes, warfarin, hip pain, back pain, glucosamine use, radiographic severe, study knee JSW, previous analgesic use, and baseline values of outcomes; Group 1 Number missing: 27, Reason: 8 adverse events, 19 other; Group 2 Number missing: 34, Reason: 8 adverse events, 26 other Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiac problems at 2 years; Group 1: 0/152, Group 2: 1/151 Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Reports only adverse events leading to withdrawal; Baseline details: Reports age, BMI, gender, presence of Heberden's nodes, knee surgery history, comorbidity, diabetes, warfarin, hip pain, back pain, glucosamine use, radiographic severe, study knee JSW, previous analgesic use, and baseline values of outcomes; Group 1 Number missing: 27, Reason: 8 adverse events, 19 other; Group 2 Number missing: 34, Reason: 8 adverse events, 26 other	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Lehmann 2005[117]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1684)
Countries and setting	Conducted in Unknown multicentre; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis meeting the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged at least 18 years with symptomatic primary knee osteoarthritis who were expected to require chronic NSAID or other analgesic therapy for at least 3 months.
Exclusion criteria	Secondary osteoarthritis and/or a history/evidence of significant diseases in the affected knee; open knee surgery (in the affected knee) within 1 year; observational arthroscopy; arthroscopic surgery or lavage within 180 days; other connective tissue disease or significant medical problems (including fibromyalgia, sarcoidosis, rheumatoid arthritis, adult juvenile chronic arthritis, peptic ulceration, GI bleeding, history of malignancy); known hypersensitivity to analgesics, antipyretics or NSAIDs; pregnant or lactating women or those of childbearing potential not using a reliable method of contraception; people with evidence of hepatic, renal or blood coagulation disorders or anaemia
Recruitment/selection of patients	People underwent a screening period of 3-7 days, during which prior NSAID therapy was washed out (with the exception of rescue paracetamol or aspirin for prophylaxis against a cardio- or cerebrovascular event). The screening period was more than or equal to 14 days in people who had been receiving piroxicam and similar agents with long half-lives (>60 hours). Following screening, people with OA pain intesnity in the target knee of at least 40mm on a 100mm visual analogue scale during the past 24 hours were eligible to enter the treatment phase. However, no increase in OA pain intensity (flare) from screening to baseline was required.
Age, gender and ethnicity	Age - Mean (SD): 62.4 (10.1). Gender (M:F): 513:1171. Ethnicity: Caucasian = 1662, Others not specified

Further population details	1. Age: Mixed (Basedo on range: 22-91). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: 4.2 (5.8) years
Indirectness of population	No indirectness
Interventions	(n=420) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 13 weeks. Concurrent medication/care: Paracetamol less than or equal to 2g/day was permitted as rescue medication throughout the study, including during the screening/washout period. Concomitant NSAID therapy was not allowed during the study, with the exception of low dose aspirin (less than or equal to 325mg/day) for prophylaxis against cardiovascular or cerebrovascular events in people considered at increased risk. Other permitted concomitant medications included chondroitin sulphate and/or glucosamine sulphate (if established stable dosage and regimen); corticosteroids (topical, ophthalmic, nasal or inhaled at usual labeled doses); histamine-2 receptor antagonists; proton pump inhibitors, antacids and cytoprotective agents (taken at usual labeled doses); and physiotherapy as prescribed by the physician (only when ongoing with a stable regimen).. Indirectness: No indirectness (n=840) Intervention 2: NSAIDs - Other. Lumiracoxib 100mg once a day with or without a loading dose. Duration 13 weeks. Concurrent medication/care: Paracetamol less than or equal to 2g/day was permitted as rescue medication throughout the study, including during the screening/washout period. Concomitant NSAID therapy was not allowed during the study, with the exception of low dose aspirin (less than or equal to 325mg/day) for prophylaxis against cardiovascular or cerebrovascular events in people considered at increased risk. Other permitted concomitant medications included chondroitin sulphate and/or glucosamine sulphate (if established stable dosage and regimen); corticosteroids (topical, ophthalmic, nasal or inhaled at usual labeled doses); histamin-2 receptor antagonists; proton pump inhibitors, antacids and cytoprotective agents (taken at usual labeled doses); and physiotherapy as prescribed by the physician (only when ongoing with a stable regimen).. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=424) Intervention 3: Placebo. Matching placebo once a day. Duration 13 weeks. Concurrent medication/care: Paracetamol less than or equal to 2g/day was permitted

	as rescue medication throughout the study, including during the screening/washout period. Concomitant NSAID therapy was not allowed during the study, with the exception of low dose aspirin (less than or equal to 325mg/day) for prophylaxis against cardiovascular or cerebrovascular events in people considered at increased risk. Other permitted concomitant medications included chondroitin sulphate and/or glucosamine sulphate (if established stable dosage and regimen); corticosteroids (topical, ophthalmic, nasal or inhaled at usual labeled doses); histamine-2 receptor antagonists; proton pump inhibitors, antacids and cytoprotective agents (taken at usual labeled doses); and physiotherapy as prescribed by the physician (only when ongoing with a stable regimen).. Indirectness: No indirectness
Funding	Study funded by industry (This work was supported by a grant from Novartis Pharma AG, Basel, Switzerland)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -3.4 (SD 3.67); n=420, Group 2: mean -2.5 (SD 4.12); n=424; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 10.2 (3.23). Baseline placebo: 9.8 (3.20).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, OA disease duration and baseline values of outcomes; Group 1 Number missing: 52, Reason: 52 discontinued. 20 adverse events, 13 unsatisfactory therapeutic effect, 8 protocol violation, 10 withdrawal of consent, 1 administrative problem.; Group 2 Number missing: 64, Reason: 64 discontinued. 16 adverse events, 31 unsatisfactory therapeutic effect, 10 protocol violation, 4 withdrawal of consent, 3 lost to follow up

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC difficulty in performing daily activities subscale at 13 weeks; Group 1: mean -10.3 (SD 11.76); n=420, Group 2: mean -8 (SD 13.29); n=424; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline celecoxib: 36.2 (10.78). Baseline placebo: 35.8 (10.84).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, OA disease duration and baseline values of outcomes; Group 1 Number missing: 52, Reason: 52 discontinued. 20 adverse events, 13 unsatisfactory therapeutic effect, 8 protocol violation, 10 withdrawal of consent, 1 administrative problem.; Group 2 Number missing: 64, Reason: 64 discontinued. 16 adverse events, 31 unsatisfactory therapeutic effect, 10 protocol violation, 4 withdrawal of consent, 3 lost to follow up

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Hypertension NOS at 13 weeks; Group 1: 8/420, Group 2: 14/424
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, OA disease

duration and baseline values of outcomes; Group 1 Number missing: 52, Reason: 52 discontinued. 20 adverse events, 13 unsatisfactory therapeutic effect, 8 protocol violation, 10 withdrawal of consent, 1 administrative problem.; Group 2 Number missing: 64, Reason: 64 discontinued. 16 adverse events, 31 unsatisfactory therapeutic effect, 10 protocol violation, 4 withdrawal of consent, 3 lost to follow up Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache NOS at 13 weeks; Group 1: 29/420, Group 2: 28/424 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, OA disease duration and baseline values of outcomes; Group 1 Number missing: 52, Reason: 52 discontinued. 20 adverse events, 13 unsatisfactory therapeutic effect, 8 protocol violation, 10 withdrawal of consent, 1 administrative problem.; Group 2 Number missing: 64, Reason: 64 discontinued. 16 adverse events, 31 unsatisfactory therapeutic effect, 10 protocol violation, 4 withdrawal of consent, 3 lost to follow up	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Leigh 1989[118]
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	1 (n=12)
Countries and setting	Conducted in United Kingdom; Setting: Inpatient
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiological evidence of osteoarthritis of the hip or knee or both and symptoms sufficient to merit admission to a hospital for rheumatic diseases for intensive physiotherapy
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with radiological evidence of osteoarthritis of the hip or knee or both and symptoms sufficient to merit admission to a hospital for rheumatic diseases for intensive physiotherapy
Exclusion criteria	Fibromyalgia syndrome; history of psychiatric, severe hepatic, renal or cardiac diseases
Recruitment/selection of patients	The study started with a 1 week period for washout of previous medications
Age, gender and ethnicity	Age - Mean (range): 63 (48-72). Gender (M:F): 12:0. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=12) Intervention 1: NSAIDs - Tenoxicam. Tenoxicam 20mg at night. Duration 3 weeks. Concurrent medication/care: Paracetamol 500mg four times a day was taken in constant dosage by all people throughout the study. No other analgesic or non-steroidal anti-inflammatory drug was permitted during the study period. Psychotropic drugs were forbidden for a six week period prior to and during the study. People were allowed to take other routine concomitant medication provided the dose remained constant.. Indirectness: No indirectness (n=12) Intervention 2: Placebo. Placebo once a day at night. Duration 3 weeks.

	Concurrent medication/care: Paracetamol 500mg four times a day was taken in constant dosage by all people throughout the study. No other analgesic or non-steroidal anti-inflammatory drug was permitted during the study period. Psychotropic drugs were forbidden for a six week period prior to and during the study. People were allowed to take other routine concomitant medication provided the dose remained constant.. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TENOXICAM versus PLACEBO Protocol outcome 1: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Light headedness at 3 weeks; Group 1: 1/12, Group 2: 0/12 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports mean age and gender of the whole cohort in the text. No additional information.; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Leung 2002[119]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=501)
Countries and setting	Conducted in Argentina, Australia, Belgium, Brazil, Canada, Chile, Colombia, Costa Rica, Germany, Guatemala, Hungary, Israel, Italy, Mexico, Peru, South Africa, United Kingdom, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (followed up by a 12 week open phase, which is not included in this analysis as it undoes randomisation to achieve this)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A diagnosis of osteoarthritis of the knee or hip, based on clinical and radiographic criteria, including joint space narrowing of the primary study joint
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women at least 40 years of age with a diagnosis of osteoarthritis of the knee or hip, based on clinical and radiographic criteria, including joint space narrowing of the primary study joint; American Rheumatism Association functional class I-III; symptoms for at least 6 months; in otherwise good general health; postmenopausal or demonstrably nongravid women; regular users (for 25 of the 30 days prior to screening) of NSAIDs, selective COX-2 inhibitors, or paracetamol, and to demonstrate a minimum level of disease activity at screening, and <80mm on the WOMAC question 1 (pain walking on a flat surface). Users of NSAIDs or selective COX-2 inhibitors were required to demonstrate worsening of pain (flare) after a prespecified washout period based on the half-life of the drug. The flare criteria were: at least 40mm and an increase of at least 15mm compared with screening values on question 1 of WOMAC questionnaire and a worsening on the investigator's global assessment of disease status by at least 1 point on a 5-point Likert scale. Prestudy paracetamol users had to demonstrate reproducible symptoms on the screening and randomisation visits: of at least 40mm pain while walking on a flat surface and patient's global assessment of disease status.
Exclusion criteria	People with a past history of coronary atherosclerotic disease with active angina or congestive heart failure were excluded as were those with uncontrolled hypertension or a history of stroke, transient ischaemic attack or hepatitis in the previous two years. People with any medical condition which, in the opinion of the investigator, could confound study results or cause undue risk to the patient (e.g. co-morbid conditions

	for which NSAIDs are contraindicated) were not allowed to participate. People using concomitant warfarin, anti-epileptics, ticlopidine, clopidogrel or digoxin were also excluded. People who had received intra-articular steroids or immunosuppressant therapy within three months, or systemic steroids, misoprostol, or sucralfate within one month prior to study entry were excluded, as were regular users (defined as >6 of the 30 days prior to randomisation) of proton pump inhibitors or histamine-2 receptor blockers.
Recruitment/selection of patients	Users of NSAIDs or selective COX-2 inhibitors were required to demonstrate worsening of pain (flare) after a prespecified washout period based on the half-life of the drug. The flare criteria were: at least 40mm and an increase of at least 15mm compared with screening values on question 1 of WOMAC questionnaire and a worsening on the investigator's global assessment of disease status by at least 1 point on a 5-point Likert scale. Prestudy paracetamol users had to demonstrate reproducible symptoms on the screening and randomisation visits: of at least 40mm pain while walking on a flat surface and patient's global assessment of disease status.
Age, gender and ethnicity	Age - Mean (SD): 63.16 (9.19). Gender (M:F): 392:109. Ethnicity: Asian = 3, Black = 9, Multiracial = 49, Other = 78, White = 362
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American Rheumatism Functional class I-III (median class II) Duration of symptoms: 6.09 (6.25) years.
Indirectness of population	No indirectness
Interventions	(n=445) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice daily and matching placebo or Etoricoxib 60mg once daily and matching placebo. Duration 12 weeks. Concurrent medication/care: People were permitted to use open-label paracetamol (up to 2600mg/day) for osteoarthritis pain of the study joint not adequately controlled by study medication. Indirectness: No indirectness Comments: Etoricoxib and naproxen groups were combined due to class effect as agreed in the protocol (n=56) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: People were permitted to use open-label paracetamol (up to 2600mg/day) for osteoarthritis pain of the study joint not adequately controlled by study medication. Indirectness: No indirectness

Funding	Study funded by industry (This research was supported by a grant from Merck & Co., Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/ETORICOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -25.54 (SD 21.44); n=445, Group 2: mean -15.33 (SD 20.5); n=56; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and 95% confidence intervals. Reported etoricoxib: -25.76 (-28.58, -22.94). Reported naproxen: -25.32 (-28.13, -22.50). Reported placebo: -15.33 (-20.70, -9.96). Calculated SD etoricoxib: 21.53. Calculated SD naproxen: 21.35. Calculated SD placebo: 20.50. Baseline etoricoxib: 64.91 (16.76). Baseline naproxen: 65.64 (17.13). Baseline placebo: 68.70 (15.67).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, primary joint, prior therapy, ARA functional class, age, body weight, osteoarthritis duration, WOMAC pain subscale, WOMAC physical function subscale and patient global assessment of disease status; Group 1 Number missing: 57, Reason: Naproxen: Lack of efficacy = 7, clinical adverse events = 24, laboratory adverse events = 0, other reasons = 6. Etoricoxib: Lack of efficacy = 7, clinical adverse events = 5, laboratory adverse events = 0, other reasons = 8. ; Group 2 Number missing: 12, Reason: Placebo: Lack of efficacy = 6, clinical adverse events = 6, laboratory adverse events = 0, other reasons = 0.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -20.81 (SD 21.33); n=445, Group 2: mean -12.46 (SD 20.39); n=56; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and 95% confidence intervals. Reported etoricoxib: -20.88 (-23.69, -18.08). Reported naproxen: -20.73 (-23.53, -17.93). Reported placebo: -12.46 (-17.80, -7.12). Calculated SD etoricoxib: 21.42. Calculated SD naproxen: 21.24. Calculated SD placebo: 20.39. Baseline etoricoxib: 64.03 (18.82). Baseline naproxen: 63.71 (18.01). Baseline placebo: 68.95 (14.38).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, primary joint, prior therapy, ARA functional class, age, body weight, osteoarthritis duration, WOMAC pain subscale, WOMAC physical function subscale and patient global assessment of disease status; Group 1 Number missing: 57, Reason: Naproxen: Lack of efficacy = 7, clinical adverse events = 24, laboratory adverse events = 0, other reasons = 6. Etoricoxib: Lack of efficacy = 7, clinical adverse events = 5, laboratory adverse events = 0, other reasons = 8. ; Group 2 Number missing: 12, Reason: Placebo: Lack of efficacy = 6, clinical adverse events = 6, laboratory adverse events = 0, other reasons = 0.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Confirmed peptic ulcer bleeds at 12 weeks; Group 1: 5/445, Group 2: 0/56
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, primary joint, prior therapy, ARA functional class, age, body weight, osteoarthritis duration, WOMAC pain subscale, WOMAC physical function subscale and patient global assessment of disease status; Group 1 Number missing: 57, Reason: Naproxen: Lack of efficacy = 7, clinical adverse events = 24, laboratory adverse events = 0, other reasons = 6. Etoricoxib: Lack of efficacy = 7, clinical adverse events = 5, laboratory adverse events = 0, other reasons = 8. ; Group 2 Number missing: 12, Reason: Placebo: Lack of efficacy = 6, clinical adverse events = 6, laboratory adverse events = 0, other reasons = 0.

- Actual outcome for Other: Gastrointestinal nuisance symptoms at 12 weeks; Group 1: 118/445, Group 2: 11/56; Comments: Including ... Etoricoxib: abdominal pain = 4, diarrhoea = 10, dyspepsia = 6, epigastric discomfort = 11, heartburn = 9, nausea = 9. Naproxen: abdominal pain = 12, diarrhoea = 7, dyspepsia = 11, epigastric discomfort = 18, heartburn = 14, nausea = 12. Placebo: abdominal pain = 1, diarrhoea = 3, dyspepsia = 1, epigastric discomfort = 3, heartburn = 4, nausea = 2 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, primary joint, prior therapy, ARA functional class, age, body weight, osteoarthritis duration, WOMAC pain subscale, WOMAC physical function subscale and patient global assessment of disease status; Group 1 Number missing: 57, Reason: Naproxen: Lack of efficacy = 7, clinical adverse events = 24, laboratory adverse events = 0, other reasons = 6. Etoricoxib: Lack of efficacy = 7, clinical adverse events = 5, laboratory adverse events = 0, other reasons = 8. ; Group 2 Number missing: 12, Reason: Placebo: Lack of efficacy = 6, clinical adverse events = 6, laboratory adverse events = 0, other reasons = 0.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Lohmander 2005[120]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=970)
Countries and setting	Conducted in Argentina, Brazil, Hungary, Mexico, Norway, Poland, South Africa, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee or hip of at least three months duration with all people having radiographic evidence and fulfilling the American College of Rheumatology global functional class I-III status
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged 40 to 75 years with symptomatic osteoarthritis of the knee or hip of at least three months' duration. All people had radiographic evidence of hip or knee osteoarthritis (qualified as American College of Rheumatology global functional class I, II or III) and were current NSAID or paracetamol (acetaminophen) users. Helicobacter pylori status was assessed by serology at screening.
Exclusion criteria	Osteoarthritis secondary to inflammatory joint disease; a diagnosis of arthritis other than osteoarthritis; a history of gastric or duodenal bleeding within 6 months, or gastric or duodenal ulcer within 3 months; NSAID hypersensitivity; history of orthostatic hypotension; endoscopic ulcers at baseline screening; people on aspirin, histamine-2 receptor antagonists, antacids, misoprostol, proton pump inhibitors, or sucralfate.
Recruitment/selection of patients	Only previous users of paracetamol or NSAIDs were permitted into the study. All people taking NSAIDs had to withdraw from the drug 7-10 days before the baseline assessment.
Age, gender and ethnicity	Age - Mean (SD): 59.3 (8.5). Gender (M:F): 264:706. Ethnicity: White = 777, Black = 33, Other = 160
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 3 months..
Indirectness of population	No indirectness

Interventions	(n=417) Intervention 1: NSAIDs - Naproxen. Naproxen 500g twice daily. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg a day, provided by the investigator, for the control of pain during the washout period. If they used paracetamol, it was requested that it be discontinued 12 hours before the baseline visit.. Indirectness: No indirectness (n=437) Intervention 2: NSAIDs - Naproxen. AZD3582 (COX-inhibiting nitric oxide donator) 750mg twice a day. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg a day, provided by the investigator, for the control of pain during the washout period. If they used paracetamol, it was requested that it be discontinued 12 hours before the baseline visit.v. Indirectness: No indirectness Comments: AZD3582 is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol (n=116) Intervention 3: Placebo. Placebo twice a day. Duration 6 weeks. Concurrent medication/care: People were allowed to take paracetamol up to 4000mg a day, provided by the investigator, for the control of pain during the washout period. If they used paracetamol, it was requested that it be discontinued 12 hours before the baseline visit.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by a grant from Astra Zeneca)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulcers at 6 weeks; Group 1: 57/417, Group 2: 0/116; Comments: Naproxen: 13.7% (57). Placebo: 0
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, race, primary study joint, weight, height, BMI, gastroduodenal erosions and ulcers, and baseline WOMAC scores; Group 1 Number missing: 23, Reason: 394 completed the trial (out of those who received at least 1 dose of the investigational product and had at least two endoscopies); Group 2 Number missing: 16, Reason: 100 completed the trial (out of those who received at least 1 dose of the investigational product and had at least two endoscopies)

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Unstable angina at 6 weeks; Group 1: 1/417, Group 2: 0/116
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Reports gender, age, race, primary study joint, weight,

height, BMI, gastroduodenal erosions and ulcers, and baseline WOMAC scores; Group 1 Number missing: 23, Reason: 394 completed the trial (out of those who received at least 1 dose of the investigational product and had at least two endoscopies); Group 2 Number missing: 16, Reason: 100 completed the trial (out of those who received at least 1 dose of the investigational product and had at least two endoscopies)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Lopes vaz 1982[121]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Portugal; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a diagnosis of unilateral osteoarthritis of the knee without major complications
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a diagnosis of unilateral osteoarthritis of the knee without major complications
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 57.8 (5.5). Gender (M:F): 10:28. Ethnicity: Not stated
Further population details	1. Age: <75 years (Based on range: 48-68). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 3.2 (2.0) years
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 250mg capsules - 2 capsules 3 times a day. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=20) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 200mg - 2 capsules three times a day. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus IBUPROFEN

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Heartburn and epigastric pain, nausea and abdominal pain at 8 weeks; Group 1: 2/18, Group 2: 3/20; Comments: Glucosamine: 1 heartburn and epigastric pain, 1 nausea. Ibuprofen: 2 heartburn and epigastric pain, 1 abdominal pain
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age and duration of illness; Group 1 Number missing: 2, Reason: 2 people did not return for follow up; Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 8 weeks; Group 1: 0/18, Group 2: 1/20
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age and duration of illness; Group 1 Number missing: 2, Reason: 2 people did not return for follow up; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Lopez sanchez 1983[122]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Spain; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographically verified degenerative joint disease of one or both hips and/or knees with pain on motion in the affected joint which had been present for at least 6 months
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Adult outpatients with definite, radiographically verified degenerative joint disease of one or both hips and/or knees. All people selected had pain on motion in the affected joint which had to have been present for at least 6 months and which had not essentially changed during the last 2 months immediately preceding entrance into the study.
Exclusion criteria	People were not accepted into the study with degenerative joint diseases secondary to metabolic bone disease or coexisting diseases predisposing to subsequent degenerative diseases (such as rheumatoid diseases, ankylosing spondylitits, gout or neuropathy); concomitant use of anticoagulants; history of allergic reactions to anthranilic derivatives, phenylbutazone or naproxen; complications resulting from long-term drug therapy, such as severe osteoporosis, chronic bone marrow suppression, renal impairment etc.; severe cardiorespiratory insufficiency; clinically significant digestive disorders, vasculitis, or peripheral neuropathy; findings suggestive of hepatorenal diseases; significant ocular pathology; or degenerative articular changes as a result of trauma; people with a haemoglobin ≤10g%, haematocrit ≤30%, WBC ≤4.800/mm³, or if serum bilirubin, SGOT or SGPT, alkaline phosphatase, or serum creatinine exceeded the normal range; females of childbearing potential
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Median (range): 52.3 (39-71). Gender (M:F): 24:16. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip and/or knee (15 hip, 24 knee, 1 both)).

Extra comments	Severity: Not stated Duration of symptoms (median [range]): 2.25 (1-12) years
Indirectness of population	No indirectness
Interventions	(n=10) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg per day (taken over 3 times per day). Duration 6 weeks. Concurrent medication/care: People were asked to not take any other analgesic or anti-inflammatory medication during the study. If, however, additional analgesic medication was required, paracetamol was prescribed.. Indirectness: No indirectness (n=10) Intervention 2: NSAIDs - Other. Meclofenamate sodium 300mg per day for 6 weeks (taken over 3 times per day). Duration 6 weeks. Concurrent medication/care: People were asked to not take any other analgesic or anti-inflammatory medication during the study. If, however, additional analgesic medication was required, paracetamol was prescribed.. Indirectness: No indirectness Comments: Meclofenamate is not licensed for use in the UK and so was not included in the analysis. However, it was reported here for completeness. (n=10) Intervention 3: NSAIDs - Other. Phenylbutazone 600mg per day for the first week, then 300mg per day for 5 weeks (taken over 3 times per day). Duration 6 weeks. Concurrent medication/care: People were asked to not take any other analgesic or anti-inflammatory medication during the study. If, however, additional analgesic medication was required, paracetamol was prescribed.. Indirectness: No indirectness Comments: Phenylbutazone is not licensed for use in the UK and so was not included in the analysis. However, it was reported here for completeness. (n=10) Intervention 4: Placebo. Placebo given three times per day. Duration 6 weeks. Concurrent medication/care: People were asked to not take any other analgesic or anti-inflammatory medication during the study. If, however, additional analgesic medication was required, paracetamol was prescribed.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Dyspepsia at 6 weeks; Group 1: 3/10, Group 2: 0/10

Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, weight, joint affected and duration of disease; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Lund 1998[123]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=513)
Countries and setting	Conducted in Belgium, Denmark, Germany, Netherlands, Sweden; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinically and radiographically confirmed diagnosis of osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People were at least 18 years of age, had a clinically and radiographically confirmed diagnosis of osteoarthritis of the knee, at least moderate pain in the affected knee (at least 35mm as assessed by the patient on a 100mm Visual Analogue Scale, 0mm = no pain, 100mm = unbearable pain), and had symptoms for at least 3 months before entry to the study
Exclusion criteria	Treatment with other NSAIDs; massage and exercise routines that were changed during the course of the study
Recruitment/selection of patients	People treated with NSAIDs underwent a washout period of between 3 and 7 days - no inclusion criteria based on response after this
Age, gender and ethnicity	Age - Mean (SD): 68.5 (11.8). Gender (M:F): 112:299. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on SD). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: 8.4 (8.4) years.
Indirectness of population	No indirectness
Interventions	(n=274) Intervention 1: NSAIDs - Meloxicam. Meloxicam 7.5mg or 15mg once daily. Duration 3 weeks. Concurrent medication/care: People receiving therapy for concomitant diseases were allowed to continue with their medication. Treatment with other NSAIDs was not allowed. People could use paracetamol as rescue medication throughout the course of the study. Massage and exercise were also permitted, provided the routines continued unchanged throughout the course of the study..

	Indirectness: No indirectness Comments: The three different doses made up three different groups in the study. These were combined for this analysis due to class effect as agreed in the protocol. (n=137) Intervention 2: Placebo. Placebo once a day. Duration 3 weeks. Concurrent medication/care: People receiving therapy for concomitant diseases were allowed to continue with their medication. Treatment with other NSAIDs was not allowed. People could use paracetamol as rescue medication throughout the course of the study. Massage and exercise were also permitted, provided the routines continued unchanged throughout the course of the study.. Indirectness: No indirectness (n=102) Intervention 3: NSAIDs - Meloxicam. Meloxicam 30mg once daily. Duration 3 weeks. Concurrent medication/care: People receiving therapy for concomitant diseases were allowed to continue with their medication. Treatment with other NSAIDs was not allowed. People could use paracetamol as rescue medication throughout the course of the study. Massage and exercise were also permitted, provided the routines continued unchanged throughout the course of the study.. Indirectness: No indirectness Comments: This group was reported incompletely with it being unclear as to what the attrition rate, baseline characteristics and effects on specific outcomes were. Meloxicam 30mg is over the recommended dose in the BNF. Due to these factors the group was not included in the analysis.
Funding	Principal author funded by industry (M. Distel and E. Bluhmki are associated with Boehringer Ingelheim GmbH, Biberach/Riss, Germany)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: MELOXICAM versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain on movement at 3 weeks; Group 1: mean -24.7 (SD 25.4); n=274, Group 2: mean -18 (SD 25.5); n=137; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Combined meloxicam groups. Reported meloxicam 7.5mg: -25.2 (24.7). Reported meloxicam 15mg: -24.1 (26.1). Reported placebo: -18.0 (25.5). Baseline meloxicam 7.5mg: 63.8 (16.3). Baseline meloxicam 15mg: 60.4 (16.3). Baseline placebo: 59.6 (16.3).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study does not report the meloxicam 30mg group results in full. Therefore, outcomes were downgraded for risk of bias due to outcome reporting bias.; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis and baseline values for pain on movement and at rest; Group 1 Number missing: 16, Reason: Reports that 7 people in the meloxicam 7.5mg group and 9 people in the meloxicam 15mg group withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.; Group 2 Number missing: 10, Reason: Reports that 10 people withdrew due to adverse events. But does not state overall

population (only the ITT population) and does not state any other reasons for attrition.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events at 3 weeks; Group 1: 35/274, Group 2: 17/137; Comments: Meloxicam 7.5mg: 12.9%. Meloxicam 15mg: 12.7%. Meloxicam 30mg: Not reported. Placebo: 12.4%.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study does not report the meloxicam 30mg group results in full. Therefore, outcomes were downgraded for risk of bias due to outcome reporting bias.; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis and baseline values for pain on movement and at rest; Group 1 Number missing: 16, Reason: Reports that 7 people in the meloxicam 7.5mg group and 9 people in the meloxicam 15mg group withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.; Group 2 Number missing: 10, Reason: Reports that 10 people withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Oedema (not otherwise specified), peripheral oedema, angina pectoris and myocardial infarction leading to withdrawal from the study at 3 weeks; Group 1: 3/274, Group 2: 2/137; Comments: Meloxicam 7.5mg: Oedema (not otherwise specified) = 0, peripheral oedema = 1, angina pectoris = 1, myocardial infarction = 0. Meloxicam 15mg: Oedema (not otherwise specified) = 1, peripheral oedema = 0, angina pectoris = 0, myocardial infarction = 0. Placebo: Oedema (not otherwise specified) = 0, peripheral oedema = 0, angina pectoris = 1, myocardial infarction = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study does not report the meloxicam 30mg group results in full. Therefore, outcomes were downgraded for risk of bias due to outcome reporting bias.; Indirectness of outcome: Serious indirectness, Comments: Only including adverse events leading to study withdrawal; Baseline details: Reports age, sex, duration of osteoarthritis and baseline values for pain on movement and at rest; Group 1 Number missing: 16, Reason: Reports that 7 people in the meloxicam 7.5mg group and 9 people in the meloxicam 15mg group withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.; Group 2 Number missing: 10, Reason: Reports that 10 people withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.

Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Micturition frequency leading to withdrawal from the study at 3 weeks; Group 1: 0/274, Group 2: 0/137; Comments: Meloxicam 7.5mg: Micturition frequency = 0, elevated alkaline phosphatase = 0, elevated gamma-GT = 1, elevated uric acid = 0, elevated urea = 0, elevated creatinine = 0, proteinuria = 0, haematuria = 0, elevated potassium = 0. Meloxicam 15mg: Micturition frequency = 1, elevated alkaline phosphatase = 1, elevated gamma-GT = 5, elevated uric acid = 3, elevated urea = 2, elevated creatinine = 2, proteinuria = 0, haematuria = 0, elevated potassium = 1. Placebo: Micturition frequency = 0, elevated alkaline phosphatase = 1, elevated gamma-GT = 3, elevated uric acid = 0, elevated urea = 0, elevated creatinine = 0, proteinuria = 0, haematuria = 0, elevated potassium = 0.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study does not report the meloxicam 30mg group results in full. Therefore, outcomes were downgraded for risk of bias due to outcome reporting bias.; Indirectness of outcome: Serious indirectness, Comments: Only including adverse events leading to study withdrawal; Baseline details: Reports age, sex, duration of osteoarthritis and baseline values for pain on movement and at rest; Group 1 Number

missing: 16, Reason: Reports that 7 people in the meloxicam 7.5mg group and 9 people in the meloxicam 15mg group withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.; Group 2 Number missing: 10, Reason: Reports that 10 people withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition. Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Dizziness, headache, vertigo and tinnitus leading to withdrawal from the study at 3 weeks; Group 1: 7/274, Group 2: 5/137; Comments: Meloxicam 7.5mg: dizziness = 0, headache = 1, vertigo = 0, tinnitus = 0. Meloxicam 15mg: dizziness = 1, headache = 1, vertigo = 2, tinnitus = 2. Placebo: dizziness = 2, headache = 2, vertigo = 1, tinnitus = 0. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study does not report the meloxicam 30mg group results in full. Therefore, outcomes were downgraded for risk of bias due to outcome reporting bias.; Indirectness of outcome: Serious indirectness, Comments: Only including adverse events leading to study withdrawal; Baseline details: Reports age, sex, duration of osteoarthritis and baseline values for pain on movement and at rest; Group 1 Number missing: 16, Reason: Reports that 7 people in the meloxicam 7.5mg group and 9 people in the meloxicam 15mg group withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.; Group 2 Number missing: 10, Reason: Reports that 10 people withdrew due to adverse events. But does not state overall population (only the ITT population) and does not state any other reasons for attrition.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Makarowski 1996[124]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=347)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis defined as the presence of knee pain with at least three of the following criteria: age older than 50 years, morning stiffness for less than 30 minutes, crepitus of the knee, bony tenderness and bony enlargement with no palpable warmth. In addition, osteoarthritis had to be confirmed radiologically on a weight-bearing roentgenogram of the knee obtained before the first dose of study medication.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults of either sex who were older than 18 years of age who weighed at least 100lb if they had osteoarthritis of the knee for at least 6 months. Osteoarthritis defined as the presence of knee pain with at least three of the following criteria: age older than 50 years, morning stiffness for less than 30 minutes, crepitus of the knee, bony tenderness and bony enlargement with no palpable warmth. In addition, osteoarthritis had to be confirmed radiologically on a weight-bearing roentgenogram of the knee obtained before the first dose of study medication. Women of childbearing potential were eligible for enrollment if they were practicing an effective method of contraception.
Exclusion criteria	A history of hypersensitivity to NSAIDs (eg, anaphylactoid or cutaneous reactions); who had any clinically significant condition that would preclude study participation; who had arthritis other than primary osteoarthritis (eg. psoriatic arthritis, syphilitic neuropathy, ochronosis, metabolic bone diseases, rheumatoid arthritis, chondrocalcinosis and gouty arthritis); who had a history of nasal polyps, bronchospasm or angioedema; who had any malignancy other than basal cell carcinoma within a year of study entry; who had a history, within 3 years of the first dose of the study medication, of document duodenal perforation, peptic perforation, or gastrointestinal haemorrhaging that required surgery; who had a history of inflammatory bowel disease or endoscopically confirmed NSAID-induced gastric ulcer within 60 days of the first dose of the study medication; or who had a positive stool test for occult blood obtained at screening or during the washout period; aspartate

	aminotransferase or alanine aminotransferase levels 1.5 times the upper limit of normal; bilirubin levels over the upper limit of normal; had abnormal laboratory values at screening that may have reflected the presence of renal or hepatic diseases; were unable to discontinue taking NSAIDs (other than study medications or 325mg for less of aspirin daily for cardiovascular prophylaxis), warfarin, sulfasalazine, corticosteroids, any analgesic other than the study medications, or tricyclic antidepressants (excluding those given for sleep at a subtherapeutic dose); had taken systemic or intra-articular corticosteroids within 30 days of the first dose of the study medication or were expected to require intra-articular corticosteroids during the study; had taken any investigational medication within 30 days of the first dose of the study medication or were scheduled to take an investigational drug other than oxaprozin during the study; had a history of drug or alcohol abuse within a year of study entry; or were anticipated to be unable to comply with the study protocol
Recruitment/selection of patients	Only including people who underwent a 14 day washout period until osteoarthritis pain became at least moderate in severity in a flare. Flare was defined as worsening of knee pain on motion or on weight bearing and worsening of both the patients' and physicians' global assessment of arthritis increased to at least a score of 2 (moderate), with a worsening from screening of at least one grade following discontinuation of NSAID therapy in the washout period
Age, gender and ethnicity	Age - Mean (range): 61.1 (26-88). Gender (M:F): 111:236. Ethnicity: 84.1% were white, 11.5% were black
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Indirectness of population	No indirectness
Interventions	(n=231) Intervention 1: NSAIDs - Other. Oxaprozin 1400mg once a day or nabumetone 1500mg once a day with double dummy placebo tablets. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The oxaprozin and nabumetone groups were pooled for analysis due to class effect as agreed in the protocol (n=116) Intervention 2: Placebo. Matching placebo once a day. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a grant from G.D. Searle & Co.. Skokie, Illinois.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXAPROZIN/NABUMETONE versus PLACEBO Protocol outcome 1: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 21/231, Group 2: 11/116; Comments: Oxaprozin: 6, Nabumetone: 15, Placebo: 11 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Makarowski 2002[125]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=467)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosed symptomatic osteoarthritis of the hip: hip pain plus at least 2 of the following criteria: Westergren erythrocyte sedimentation rate <20mm/hour, radiographic femoral or acetabular osteophytes, or radiographic joint space narrowing
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a diagnosis of symptomatic osteoarthritis of the hip. People also had to have had a baseline Patient's Assessment of Arthritis Pain-VAS score of at least 40mm and both Patient's and Physician's Global Assessment of Arthritis of poor or very poor
Exclusion criteria	People with inflammatory arthritis, gout, pseudogout, Paget's disease or any chronic pain syndrome that interfered with assessment of the Index Hip; people with osteoarthritis of the knee ipsilateral to the index hip; symptomatic trochanteric bursitis or acute joint trauma of the index hip; people who had hip arthroscopy within the previous year; had complete loss of articular cartilage on weight-bearing X-ray of the index hip; people with active GI disease, GI ulceration within 30 days of study medication, or significant bleeding disorder.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.3 (11.8). Gender (M:F): Not reported. Ethnicity: Caucasian = 442, Black = 19, Hispanic = 6
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 6.1 (6.7) years
Indirectness of population	No indirectness
Interventions	(n=118) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day. Duration 12 weeks. Concurrent medication/care: People enrolled into the study discontinued

	regular pain medication at least 2 days before the baseline assessment. Oxaprozin, piroxicam and full dose aspirin (>325mg/day) were discontinued at least 4 days before the baseline arthritis assessment. Indirectness: No indirectness (n=231) Intervention 2: NSAIDs - Other. Valdecoxib 5mg or 10mg once a day. Duration 12 weeks. Concurrent medication/care: People enrolled into the study discontinued regular pain medication at least 2 days before the baseline assessment. Oxaprozin, piroxicam and full dose aspirin (>325mg/day) were discontinued at least 4 days before the baseline arthritis assessment. Indirectness: No indirectness Comments: Valdecoxib is not licensed for use in the UK and so was not included in this analysis as agreed in the protocol. It was reported here for completeness. (n=118) Intervention 3: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: People enrolled into the study discontinued regular pain medication at least 2 days before the baseline assessment. Oxaprozin, piroxicam and full dose aspirin (>325mg/day) were discontinued at least 4 days before the baseline arthritis assessment. Indirectness: No indirectness
Funding	Study funded by industry (Sponsored by Pharmacia Corporation and Pfizer Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hip: WOMAC pain subscale at 12 weeks; Group 1: mean -2.94 (SD 4.99); n=118, Group 2: mean -1.25 (SD 4.99); n=117; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean least square changes and p-values. SD calculated from this. Reported naproxen: -2.94. Reported placebo: -1.25. Reported P = ≤0.01. Calculated SE = 0.65. Baseline naproxen: 10.5. Baseline placebo: 10.8.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, ethnicity, history of GI bleeding, gastroduodenal ulcers, patient and physician's global assessment of arthritis, and baseline values for outcomes; Group 1 Number missing: 47, Reason: 15 adverse events, 24 treatment failure, 2 lost to follow-up, 2 pre-existing violation, 4 noncompliance; Group 2 Number missing: 69, Reason: 7 adverse events, 51 treatment failure, 1 lost to follow up, 3 pre-existing violation, 7 noncompliance

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hip: WOMAC physical function subscale at 12 weeks; Group 1: mean -9.63 (SD 14.8); n=118, Group 2: mean -3.18 (SD 14.8); n=117; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean least square changes and p-values. SD calculated from this. Reported naproxen: -9.63. Reported placebo: -3.18. Reported P = <0.001. Calculated SE = 1.94. Baseline naproxen: 36.6. Baseline placebo: 37.1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, ethnicity, history of GI

bleeding, gastroduodenal ulcers, patient and physician's global assessment of arthritis, and baseline values for outcomes; Group 1 Number missing: 47, Reason: 15 adverse events, 24 treatment failure, 2 lost to follow-up, 2 pre-existing violation, 4 noncompliance; Group 2 Number missing: 69, Reason: 7 adverse events, 51 treatment failure, 1 lost to follow up, 3 pre-existing violation, 7 noncompliance Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Hip: Withdrawal due to gastrointestinal related adverse events including abdominal pain, constipation, diarrhoea, dyspepsia, nausea and withdrawals due to diverticulitis, gastroesophageal reflux, haemorrhoid bleeding at 12 weeks; Group 1: 13/118, Group 2: 2/117; Comments: Including ... Naproxen: Abdominal pain = 12, constipation = 8, diarrhoea = 6, dyspepsia = 11, nausea = 10, diverticulitis = 1, gastroesophageal reflux = 1, haemorrhoid bleeding = 0. Placebo: Abdominal pain = 5, constipation = 1, diarrhoea = 2, dyspepsia = 6, nausea = 4, diverticulitis = 0, gastroesophageal reflux = 0, haemorrhoid bleeding = 0. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal due to adverse events; Baseline details: Reports age, weight, ethnicity, history of GI bleeding, gastroduodenal ulcers, patient and physician's global assessment of arthritis, and baseline values for outcomes; Group 1 Number missing: 47, Reason: 15 adverse events, 24 treatment failure, 2 lost to follow-up, 2 pre-existing violation, 4 noncompliance; Group 2 Number missing: 69, Reason: 7 adverse events, 51 treatment failure, 1 lost to follow up, 3 pre-existing violation, 7 noncompliance Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hip: Headache at 12 weeks; Group 1: 4/118, Group 2: 14/117 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, ethnicity, history of GI bleeding, gastroduodenal ulcers, patient and physician's global assessment of arthritis, and baseline values for outcomes; Group 1 Number missing: 47, Reason: 15 adverse events, 24 treatment failure, 2 lost to follow-up, 2 pre-existing violation, 4 noncompliance; Group 2 Number missing: 69, Reason: 7 adverse events, 51 treatment failure, 1 lost to follow up, 3 pre-existing violation, 7 noncompliance	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Malonne 2004[126]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=230)
Countries and setting	Conducted in Belgium; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip or knee as defined by the European League Against Rheumatism criteria if they had hip or knee pain that rated ≥35mm on the 100-mm Huskisson horizontal visual analogue scale (VAS) (scale from 0 = no pain, 100 = worst pain) and functional discomfort rated ≥4 on the Lequesne functional discomfort index (total score from 0 = absence of pain to 20 = most intense pain). Symptomatology must have been evolving for ≥6 months.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 45 to 80 years with osteoarthritis of the hip or knee (as defined by the European League Against Rheumatism criteria). Symptomatology must have been evolving for ≥6 months, requiring regular treatment with analgesics or NSAIDs for ≥1 month.
Exclusion criteria	People with ≥1 of the following: secondary osteoarthritis; genu varum >10 degrees; rapidly evolving destructive osteoarthritis; isolated or predominantly patellofemoral osteoarthritis; renal or hepatic dysfunction; or serious concomitant medical illness; if they had been receiving corticosteroid injections within 1 month before the study; had undergone irradiation of the synovial tissue in the preceding 3 months; had received orally administered corticosteroids within 1 month; analgesics within 24 hours; NSAIDs within 48 hours; oxicams within 5 days of study inclusion
Recruitment/selection of patients	Multicenter, parallel-group, 14 day trial
Age, gender and ethnicity	Age - Mean (SD): 66.7 (8.3). Gender (M:F): 63:167. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on inclusion criteria range). 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee (majority knee)).
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 5.7 (5.1) years.
Indirectness of population	No indirectness

Interventions	(n=111) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol LP 200mg once daily. Duration 14 days. Concurrent medication/care: During the second week of the study, concomitant use of paracetamol ≤2g/day (500mg capsules) was allowed as rescue analgesia for uncontrolled pain. Indirectness: No indirectness (n=119) Intervention 2: Placebo. Matching placebo once daily. Duration 14 days. Concurrent medication/care: During the second week of the study, concomitant use of paracetamol ≤2g/day (500mg capsules) was allowed as rescue analgesia for uncontrolled pain. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain (visual analogue scale) at 14 days; Group 1: mean 2.43 (SD 2.35); n=85, Group 2: mean 1.55 (SD 2.35); n=112; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reports means and a p-value. SD calculated from this. Reported tramadol: 2.43. Reported placebo: 1.55. Reported P-value = 0.010.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, BMI, most painful site, time since diagnosis, time from the start of the current episode of pain, duration of pain, number of days of pain for those with pain for less than 1 month. Does not report baseline pain value.; Group 1 Number missing: 26, Reason: No visual analogue score at day 14 = 26; Group 2 Number missing: 7, Reason: No visual analogue score at day 14 = 7

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	March 1993[128] (March 1994[127])
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	25 (n of 1 trials) (n=25)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks x 3 cycles (2 weeks x 3 cycles for each drug)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of osteoarthritis with daily pain
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	A clinical diagnosis of osteoarthritis with daily pain, no known contraindication to non-steroidal anti-inflammatory drugs, and no corticosteroid injection in the previous 4 weeks.
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Median (range): 64 (38-85). Gender (M:F): 5:20. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Unclear site).
Extra comments	Severity: Not stated Duration of symptoms: Median 8 years (range 1-30 years). N of 1 trials
Indirectness of population	No indirectness
Interventions	(n=25) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 1 gram twice daily. Duration Two weeks x 3 (Three four week cycles in total for the study). Concurrent medication/care: Paracetamol was available as an escape analgesic up to a maximum of 2 grams/day. Indirectness: Serious indirectness; Indirectness comment: Paracetamol was the intervention and available as the rescue medication (n=25) Intervention 2: NSAIDs - Diclofenac. Diclofenac 50mg twice daily. Duration Two weeks x 3 (Three four week cycles in total for the study). Concurrent medication/care: Paracetamol was available as an escape analgesic up to a maximum of 2 grams/day. Indirectness: Serious indirectness; Indirectness comment:

	Paracetamol was the comparator and the rescue medication
Funding	Study funded by industry (Financial support was provided by the Northern Sydney Area Health Service and Sterling Phaermaceuticals Pty Ltd Australia. Ciba-Geigy Australia supplied the diclofenac sodium.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus DICLOFENAC

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain (visual analogue scale) at 2 weeks; Group 1: mean 25.5 (SD 22.7); n=15, Group 2: mean 18.6 (SD 20.2); n=15; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports the mean value for each participant and the p-value for each participant. Each participant combined together for analysis. Case 2: paracetamol = 9.3, NSAID = 7.5, p = 0.23. Case 3: paracetamol = 54.3, NSAID = 50.1, p = 0.1. Case 5: paracetamol = 14.1, NSAID = 0.4, p = 0.19. Case 6: paracetamol = 62.1, NSAID = 62.5, p = 0.63. Case 7: paracetamol = 34.3, NSAID = 38.4, p = 0.63. Case 8: paracetamol = 7.2, NSAID = 4.8, p = 0.14. Case 9: paracetamol = 23.8, NSAID = 5.8, p = 0.001. Case 11: paracetamol = 5.8, NSAID = 4.3, p = 0.36. Case 12: paracetamol = 40.5, NSAID = 35.0, p = 0.33. Case 13: paracetamol = 9.7, NSAID = 12.5, p = 0.63. Case 15: paracetamol = 65.9, NSAID = 36.4, p = 0.01. Case 19: paracetamol = 2.1, NSAID = 0.8, p = 0.11. Case 20: paracetamol = 7.2, NSAID = 7.3, p = 0.55. Case 21: paracetamol = 43.3, NSAID = 9.5, p = 0.003. Case 23: paracetamol = 2.7, NSAID = 3.1, p = 0.67.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Comparing against themselves; Group 1 Number missing: 10, Reason: 5 withdrew early having made a decision after two cycles, 5 dropped out very early (2 because of intolerance of the NSAID, 2 because of intercurrent illness, 1 because of non-compliance).; Group 2 Number missing: 10, Reason: 5 withdrew early having made a decision after two cycles, 5 dropped out very early (2 because of intolerance of the NSAID, 2 because of intercurrent illness, 1 because of non-compliance).

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Matsumoto 2005[129] (Mcilwain 2005[130])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=489)
Countries and setting	Conducted in Unknown multicentre; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis with the presence of typical knee or hip joint symptoms and signs and radiographic evidence of osteoarthritis, with a minimum of grade 2 in the index joint using the Kellgren-Lawrence scale.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee or hip who have taken either paracetamol, a conventional NSAID, a COX-2 inhibitor, or an opioid analgesic for at least 75 of 90 days before the screening visit and must have had a suboptimal response to these agents. Age >40 years, use of a medically acceptable form of contraception or abstinence in women of childbearing potential, and a negative serum pregnancy test 7 days before first dose of study medication. Eligible people entered a 2- to 7-day washout period during which all analgesic medications were discontinued. People were randomised when pain in the index joint reached 40mm on a 0-100mm visual analogue scale.
Exclusion criteria	People with inflammatory arthritis, gout, Paget's disease, chronic pain syndrome, or fibromyalgia; people requiring knee or hip arthroplasty within 2 months of screening anticipating any need for surgical procedures on the index joint during the study; weight <100 pounds; difficulty swallowing capsules or tablets; prior history of substance or alcohol abuse; corticosteroid or investigational drug use within 1 month of first study treatment; prior history of intolerance to opioids
Recruitment/selection of patients	Eligible people entered a 2- to 7-day washout period during which all analgesic medications were discontinued. People were randomised when pain in the index joint reached 40mm on a 0-100mm visual analogue scale.
Age, gender and ethnicity	Age - Other: Mean (SE): 62.3 (0.98). Gender (M:F): 192:297. Ethnicity: White = 422, Black = 51, Hispanic = 12, Other = 3
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).

Extra comments	Severity: Minimum Kellgren Lawrence grade 2 Duration of disease: Not explicitly stated. 146 were <5 years. 342 were more than or equal to 5 years.
Indirectness of population	No indirectness
Interventions	(n=125) Intervention 1: Strong opioids (oral) - Oxycodone. Oxycodone CR 20mg given every 12 hours. The people in the oxycodone 20mg group received oxycodone 10mg for one week and then were uptitrated to 20mg.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=124) Intervention 2: Placebo. Placebo every 12 hours. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=240) Intervention 3: Strong opioids (oral) - Other. Oxymorphone ER 40mg or oxymorphone 20mg ER given every 12 hours. The people in the oxymorphone 40mg group received oxymorphone 20mg ER for one week and then were uptitrated to 40mg. The people in the oxymorphone 20mg group received oxymorphone 10mg ER for one week and then were uptitrated to 20mg.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: This group is not included in the final analysis as oxymorphone is not licensed for use in the United Kingdom
Funding	Study funded by industry (This study was supported by Endo Pharmaceuticals Inc., Chadds Ford, Pennsylvania, and Penwest Phaemceuticals Co., Danbury, Connecticut)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYCODONE versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months

- Actual outcome for Other: SF-36 physical component at 4 weeks; Group 1: mean 4 (SD 8.9); n=125, Group 2: mean 1.8 (SD 7.8); n=124; SF-36 physical component 0-100 Top=High is good outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported oxycodone: 4.0 (0.8). Reported placebo: 1.8 (0.7). Calculated SD oxycodone: 8.9. Calculated SD placebo: 7.8.

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, weight, height, body mass index, duration of disease and index joint. Does not report baseline values for outcomes.; Group 1 Number missing: 50, Reason: Oxycodone: 50 discontinued. 13 insufficient therapeutic response. 31 nonserious adverse event. 2 noncompliance. 3 patient request. 1 other.; Group 2 Number missing: 46, Reason: 46 discontinuation. 24 insufficient therapeutic response, 6 nonserious AE, 1 noncompliance, 2 patient request, 1 investigator decision, 1 lost to follow up, 1 other.

- Actual outcome for Other: SF-36 mental component at 4 weeks; Group 1: mean -0.8 (SD 10.1); n=125, Group 2: mean 2.2 (SD 10); n=124; SF-36 mental component 0-100 Top=High is good outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported oxycodone: -0.8 (0.9). Reported placebo: 2.2 (0.9). Calculated SD oxycodone: 10.1. Calculated SD placebo: 10.0. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, weight, height, body mass index, duration of disease and index joint. Does not report baseline values for outcomes.; Group 1 Number missing: 50, Reason: Oxycodone: 50 discontinued. 13 insufficient therapeutic response. 31 nonserious adverse event. 2 noncompliance. 3 patient request. 1 other.; Group 2 Number missing: 46, Reason: 46 discontinuation. 24 insufficient therapeutic response, 6 nonserious AE, 1 noncompliance, 2 patient request, 1 investigator decision, 1 lost to follow up, 1 other. Protocol outcome 2: Pain reduction at ≤3- or >3- months - Actual outcome for Other: Patient's global assessment (Visual analogue scale) at 4 weeks; Group 1: mean -25.4 (SD 31.3); n=125, Group 2: mean -19.5 (SD 30.1); n=124; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported oxycodone: -25.4 (2.8). Reported placebo: -19.5 (2.7). Calculated SD oxycodone: 31.3. Calculated SD placebo: 30.1. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, weight, height, body mass index, duration of disease and index joint. Does not report baseline values for outcomes.; Group 1 Number missing: 50, Reason: Oxycodone: 50 discontinued. 13 insufficient therapeutic response. 31 nonserious adverse event. 2 noncompliance. 3 patient request. 1 other.; Group 2 Number missing: 46, Reason: 46 discontinuation. 24 insufficient therapeutic response, 6 nonserious AE, 1 noncompliance, 2 patient request, 1 investigator decision, 1 lost to follow up, 1 other.	
Protocol outcomes not reported by the study	Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Mckenna 2001[132]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=182)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis as defined by the American College of Rheumatology criteria (unclear if imaging was involved)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults aged 40 years or older if they met the classification of knee osteoarthritis based on the American College of Rheumatology criteria, were in a functional capacity classification of I-III at screening, and had symptoms of osteoarthritis that were defined as worsening of the signs and symptoms of the disease following discontinuation of treatment with NSAIDs or other analgesic medications, or defined by other criteria if the person was not receiving treatment (flare criteria)
Exclusion criteria	Women of childbearing potential who did not use adequate contraception and had positive serum or urinary pregnancy tests; presence of a significant active malignancy within 5 years; inflammatory arthritis or acute joint trauma of the knee; active gastrointestinal, renal or hepatic disease; a coagulation defect; clinically significant abnormal screening laboratory values; known hypersensitivity to COX-2 specific inhibitors, sulphonamides or NSAIDs; if surgery or an invasive procedures was planned during the study; if they had been treated with a corticosteroid within 8 weeks, intra-articular hyaluronic acid within 6 months or another investigational medication within 30 days; if they had been diagnosed with or treated for oesophageal, gastric, pyloric channel, or duodenal ulcer within 30 days
Recruitment/selection of patients	Washout period for flares before admittance into the trial. The washout period lasted 2 days. If a person had a flare then they were admitted to the study.
Age, gender and ethnicity	Age - Mean (SD): 62.2 (10.3). Gender (M:F): 53:129. Ethnicity: 81% Caucasian. Other ethnicities not reported.
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity classification I-III Duration of symptoms: 10.9 (9.5) years. Flare criteria for the study - if three of the four

	were present in people previously treated with NSAIDs or analgesics: absolute score of at least 40mm for the person's assessment of arthritis pain on a visual analogue scale; 1 or more grade increase since screening in physician's global assessment of arthritis; 1 or more grade increase since screening in the person's global assessment of arthritis; 2 or more point increase since screening in the osteoarthritis severity index
Indirectness of population	No indirectness
Interventions	(n=63) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day for 6 weeks. Duration 6 weeks. Concurrent medication/care: Paracetamol up to 2g per day was permitted as rescue analgesia during the washout period but was prohibited for at least 48 hours prior to the person's return for the baseline visits. During the treatment phase, no other analgesic medications were allowed except for occasional paracetamol for non-arthritic pain and low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis. Anticoagulant, antirheumatic and antiulcer medications were prohibited, although occasional antacid use was allowed. Indirectness: No indirectness (n=59) Intervention 2: NSAIDs - Other. Rofecoxib 25mg once a day. Duration 6 weeks. Concurrent medication/care: Paracetamol up to 2g per day was permitted as rescue analgesia during the washout period but was prohibited for at least 48 hours prior to the person's return for the baseline visits. During the treatment phase, no other analgesic medications were allowed except for occasional paracetamol for non-arthritic pain and low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis. Anticoagulant, antirheumatic and antiulcer medications were prohibited, although occasional antacid use was allowed. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK so will not be included in the analysis, but is reported here for completeness. (n=60) Intervention 3: Placebo. Matching placebo once a day. Duration 6 weeks. Concurrent medication/care: Paracetamol up to 2g per day was permitted as rescue analgesia during the washout period but was prohibited for at least 48 hours prior to the person's return for the baseline visits. During the treatment phase, no other analgesic medications were allowed except for occasional paracetamol for non-arthritic pain and low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis. Anticoagulant, antirheumatic and antiulcer medications were prohibited, although occasional antacid use was allowed. Indirectness: No indirectness

Funding	Study funded by industry (The study was supported by Pharmacia and Pfizer, Incorporated)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Visual analogue scale for pain on walking at 6 weeks; Group 1: mean -38 (SD 35); n=63, Group 2: mean -25 (SD 35); n=60; Visual analogue scale (pain on walking) 0-100 Top=High is poor outcome; Comments: Reports least square mean change from baseline and p-value. Standard deviation calculated from this. P-value = 0.008. Baseline celecoxib: 71.4 (18.8). Baseline placebo: 72.9 (16.7).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, ethnicity, duration of osteoarthritis, previous gastrointestinal pathology, prescreening osteoarthritis treatment, baseline assessment scores; Group 1 Number missing: 14, Reason: 1 lost to follow up, 1 pre-existing violation, 1 protocol noncompliance, 5 treatment failure, 4 adverse event, 2 other; Group 2 Number missing: 16, Reason: 2 lost to follow up, 1 protocol noncompliance, 12 treatment failure, 1 adverse event

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal tract adverse events in total - including diarrhoea, dyspepsia and abdominal pain at 6 weeks; Group 1: 7/63, Group 2: 6/60; Comments: Celecoxib: Total = 7. Diarrhoea = 3, dyspepsia = 2, abdominal pain = 1. Placebo: Total = 6. Diarrhoea = 0. Dyspesia = 4. Abdominal pain = 0. (?unsure where the other 2 people came from).
Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, ethnicity, duration of osteoarthritis, previous gastrointestinal pathology, prescreening osteoarthritis treatment, baseline assessment scores; Group 1 Number missing: 14, Reason: 1 lost to follow up, 1 pre-existing violation, 1 protocol noncompliance, 5 treatment failure, 4 adverse event, 2 other; Group 2 Number missing: 16, Reason: 2 lost to follow up, 1 protocol noncompliance, 12 treatment failure, 1 adverse event

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 6 weeks; Group 1: 10/63, Group 2: 10/60
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, ethnicity, duration of osteoarthritis, previous gastrointestinal pathology, prescreening osteoarthritis treatment, baseline assessment scores; Group 1 Number missing: 14, Reason: 1 lost to follow up, 1 pre-existing violation, 1 protocol noncompliance, 5 treatment failure, 4 adverse event, 2 other; Group 2 Number missing: 16, Reason: 2 lost to follow up, 1 protocol noncompliance, 12 treatment failure, 1 adverse event

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-

	months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months

Study	Mckenna 2001[131]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=600)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee confirmed according to the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee confirmed according to the American College of Rheumatology criteria. This was required to be symptomatic as evidenced by a defined worsening of the signs and symptoms of the disease following discontinuation of treatment with NSAIDs or other analgesic medications
Exclusion criteria	People with active gastrointestinal disease, chronic or acute renal or hepatic disease
Recruitment/selection of patients	Conducted over 54 institutions across the United States
Age, gender and ethnicity	Age - Mean (range): 61.7 (29-88). Gender (M:F): 208:392. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Not stated / Unclear (Based on American College of Rheumatology - depends on whether this was the radiographic criteria). 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean): 8.57 years
Indirectness of population	No indirectness
Interventions	(n=400) Intervention 1: NSAIDs - Diclofenac. Diclofenac 50mg three times a day and Celecoxib 100mg twice a day. Duration 6 weeks. Concurrent medication/care: People were not permitted to receive concomitant corticosteroids, NSAIDs or intra-articular injections of hyaluronic acid. People were allowed to continue using aspirin for non-arthritis related indications if the dose was stable. Indirectness: No indirectness (n=200) Intervention 2: Placebo. Placebo. Duration 6 weeks. Concurrent medication/care: People were not permitted to receive concomitant corticosteroids, NSAIDs or intra-articular injections of hyaluronic acid. People were allowed to

	continue using aspirin for non-arthritis related indications if the dose was stable. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by Pharmacia Corporation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain at 6 weeks; Group 1: mean -4.15 (SD 4.16); n=398, Group 2: mean -2.4 (SD 4.2); n=200; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 10.6 (3.1). Baseline diclofenac: 10.7 (3.1). Baseline placebo: 10.7 (3.3).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean disease duration, index knee side, prior NSAID usage, prior NSAID gastrointestinal intolerance, prior gastrointestinal bleeds, prior gastroduodenal ulceration, cardiovascular disease prevalence, low dose aspirin usage and baseline values of outcomes; Group 1 Number missing: 79, Reason: The majority of discontinuations were due to treatment failure or adverse events, the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups; Group 2 Number missing: 71, Reason: The majority of discontinuations were due to treatment failure or adverse events, the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function at 6 weeks; Group 1: mean -14.15 (SD 13.34); n=398, Group 2: mean -12.13 (SD 13.44); n=200; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline celecoxib: 37.4 (10.4). Baseline diclofenac: 37.5 (10.4). Baseline placebo: 37.4 (11.9).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean disease duration, index knee side, prior NSAID usage, prior NSAID gastrointestinal intolerance, prior gastrointestinal bleeds, prior gastroduodenal ulceration, cardiovascular disease prevalence, low dose aspirin usage and baseline values of outcomes; Group 1 Number missing: 79, Reason: The majority of discontinuations were due to treatment failure or adverse events, the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups; Group 2 Number missing: 71, Reason: The majority of discontinuations were due to treatment failure or adverse events, the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Peripheral oedema at 6 weeks; Group 1: 15/398, Group 2: 1/200
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean disease duration, index knee side, prior NSAID usage, prior NSAID gastrointestinal intolerance, prior gastrointestinal bleeds, prior gastroduodenal ulceration, cardiovascular disease prevalence, low dose aspirin usage and baseline values of outcomes; Group 1 Number missing: 79, Reason: The majority of discontinuations were

due to treatment failure or adverse events (not possible to know how many), the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups; Group 2 Number missing: 71, Reason: The majority of discontinuations were due to treatment failure or adverse events (not possible to know how many), the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: ALT increased at 6 weeks; Group 1: 6/398, Group 2: 1/200 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean disease duration, index knee side, prior NSAID usage, prior NSAID gastrointestinal intolerance, prior gastrointestinal bleeds, prior gastroduodenal ulceration, cardiovascular disease prevalence, low dose aspirin usage and baseline values of outcomes; Group 1 Number missing: 79, Reason: The majority of discontinuations were due to treatment failure or adverse events (not possible to know how many), the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups; Group 2 Number missing: 71, Reason: The majority of discontinuations were due to treatment failure or adverse events (not possible to know how many), the other reasons (overall numbers) were loss to follow up (2), pre-existing protocol violation (6) and protocol non-compliance (22), which were evenly distributed across the treatment groups	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Melo gomes 1993[133]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=643)
Countries and setting	Conducted in Unknown multicentre; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Documented radiographic evidence (joint space narrowing, subchondral bony sclerosis; bone cysts, or gross deformity and subluxation and/or loose bodies) and symptomatic evidence of osteoarthritis of the hip and/or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People of the legal age of consent with documented radiographic evidence (joint space narrowing, subchondral bony sclerosis; bone cysts, or gross deformity and subluxation and/or loose bodies) and symptomatic evidence of osteoarthritis of the hip and/or knee of at least 3 months duration; had a functional capacity classification of I-III; had physician and patient global assessment of arthritis that were rated no better than 'fair'; were experiencing joint pain, and required continuous nonsteroid anti-inflammatory drug therapy for the duration of the study
Exclusion criteria	Any acute joint trauma at the site of osteoarthritis; chronic or acute renal or hepatic disorders; significant upper gastrointestinal mucosal damage (>10 erosions in the stomach; >10 erosions in the duodenum; or oesophageal, gastric, pyloric channel, or duodenal ulcer); any active gastrointestinal disease; use of any NSAID during the 10 days or any analgesic (other than paracetamol) during the 2 days before the initial (baseline) arthritis assessments; or known hypersensitivity to any nonsteroidal anti-inflammatory drug or any prostaglandin.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 63.6 (26-89). Gender (M:F): 155:488. Ethnicity: White = 518, Black = 58, Other = 67
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip and/or knee).

Extra comments	Severity: not stated Duration of symptoms: 0-15+ years (median 1-9.9 years)
Indirectness of population	No indirectness
Interventions	(n=216) Intervention 1: NSAID and gastroprotection - NSAID and misoprostol . Diclofenac sodium 50mg/misoprostol 200 micrograms twice daily with matching placebo. All people took one tablet and one capsule in the morning and one tablet and one capsule in the evening. Duration 4 weeks. Concurrent medication/care: Paracetamol was permitted. Indirectness: No indirectness (n=427) Intervention 2: NSAIDs - Naproxen. Either piroxicam 10mg twice daily with matching placebo or naproxen 375mg twice daily with matching placebo. All people took one tablet and one capsule in the morning and one tablet and one capsule in the evening. Duration 4 weeks. Concurrent medication/care: Paracetamol was permitted. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NSAID AND MISOPROSTOL versus PIROXICAM AND NAPROXEN Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Gastroduodenal, gastric and duodenal ulcers at 4 weeks; Group 1: 3/216, Group 2: 38/427; Comments: Including ... Diclofenac/misoprostol: Gastroduodenal ulcers = 3, gastric ulcers = 3, duodenal ulcer = 0. Piroxicam: Gastroduodenal ulcers = 21, gastric ulcers = 14, duodenal ulcers = 10. Naproxen: Gastroduodenal ulcers = 17, gastric ulcers = 15, duodenal ulcers = 3. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, ethnicity, height, weight and duration of disease; Group 1 Number missing: 16, Reason: 16 did not have a final endoscopy; Group 2 Number missing: 25, Reason: Piroxicam: 13 did not have a final endoscopy. Naproxen: 12 did not have a final endoscopy	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Miceli-richard 2004[134]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=779)
Countries and setting	Conducted in France; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis as per the Lequesne criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Symptomatic osteoarthritis of the knee for at least 3 months with a global pain intensity of the knee during physical activities for the past 24 hours of at least 30mm on a 100mm visual analogue scale
Exclusion criteria	People with a prosthesis or recent (less than 1 year) surgery of the studied knee; history of allergy to paracetamol; history of hepatitis; severe hepatic or kidney failure; steroids during the past 4 weeks; current treatment with enzymatic inductors or inhibitors; any prior history of asthma or allergy potentially requiring concomitant treatment during the study; pregnancy, lactation, or inefficacious contraception; a history of drug abuse or alcoholism
Recruitment/selection of patients	Recruited from 200 French sites
Age, gender and ethnicity	Age - Mean (SD): 70 (11). Gender (M:F): 196:583. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on SD). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: 46 (47) months
Indirectness of population	No indirectness
Interventions	(n=405) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 1 gram of oral paracetamol, four times a day with 4-6 hours between doses. Duration 6 weeks. Concurrent medication/care: People who had an intraarticular injection in the previous 4 weeks were excluded. Concomitant treatments, such as long acting osteoarthritis drugs, psychotropic or myorelaxing drugs, vitamins or minerals, had to be given at stable doses for at least 3 weeks before inclusion and during the study. Rescue drugs

	for osteoarthritis, such as oral or injectable analgesics (including paracetamol), NSAIDs and intra-articular drugs were prohibited during the whole study and patients were interviewed at each visit about their concomitant osteoarthritis drug treatment.. Indirectness: No indirectness (n=374) Intervention 2: Placebo. Matched placebo, four times a day, with 4-6 hours between intakes. Duration 6 weeks. Concurrent medication/care: People who had an intraarticular injection in the previous 4 weeks were excluded. Concomitant treatments, such as long acting osteoarthritis drugs, psychotropic or myorelaxing drugs, vitamins or minerals, had to be given at stable doses for at least 3 weeks before inclusion and during the study. Rescue drugs for osteoarthritis, such as oral or injectable analgesics (including paracetamol), NSAIDs and intra-articular drugs were prohibited during the whole study and patients were interviewed at each visit about their concomitant osteoarthritis drug treatment.. Indirectness: No indirectness
Funding	Other author(s) funded by industry (M Le Bars - International Medical Organisation/Europe, Bristol-Myers Squibb-UPSA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain visual analogue scale at 6 weeks; Group 1: mean 23 (SD 27); n=405, Group 2: mean 23 (SD 26); n=374; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline paracetamol: 66.7 (18). Baseline placebo: 69.0 (17).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, disease duration, radiological score, clinical symptoms at baseline, and prior osteoarthritis treatments; Group 1 Number missing: 107, Reason: 57 treatment failure, 36 safety, 0 lost to follow up, 15 other; Group 2 Number missing: 113, Reason: 67 treatment failure, 29 safety, 0 lost to follow up, 17 other (study reports 112)

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 6 weeks; Group 1: mean 12 (SD 17); n=405, Group 2: mean 12 (SD 16); n=374; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline paracetamol: 54 (15). Baseline placebo: 54 (15).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, disease duration, radiological score, clinical symptoms at baseline, and prior osteoarthritis treatments; Group 1 Number missing: 107, Reason: 57 treatment failure, 36 safety, 0 lost to follow up, 15 other; Group 2 Number missing: 113, Reason: 67 treatment failure, 29 safety, 0 lost to follow up, 17 other (study reports 112 in total)

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal system disorders at 6 weeks; Group 1: 46/405, Group 2: 42/374

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, disease duration, radiological score, clinical symptoms at baseline, and prior osteoarthritis treatments; Group 1 Number missing: 107, Reason: 57 treatment failure, 36 safety, 0 lost to follow up, 15 other; Group 2 Number missing: 113, Reason: 67 treatment failure, 29 safety, 0 lost to follow up, 17 other (study reports 112 in total) Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Central and peripheral nervous system disorders at 6 weeks; Group 1: 7/405, Group 2: 6/374 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Includes peripheral nervous system disorders; Baseline details: Reports gender, age, BMI, disease duration, radiological score, clinical symptoms at baseline, and prior osteoarthritis treatments; Group 1 Number missing: 107, Reason: 57 treatment failure, 36 safety, 0 lost to follow up, 15 other; Group 2 Number missing: 113, Reason: 67 treatment failure, 29 safety, 0 lost to follow up, 17 other (study reports 112 in total)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Moss 2017[135]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=86)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Painful knee osteoarthritis (visual analogue scale ≥3/10) using the American College of Rheumatology classification
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with painful knee osteoarthritis (visual analogue scale ≥3/10) using the American College of Rheumatology classification. If both knees showed osteoarthritis, the knee with the highest pain was designated the index knee.
Exclusion criteria	History of systemic inflammatory condition; neurological disorders affecting sensory, motor or cognitive function; recent lower limb injury or surgery; history of other chronic pain disorders
Recruitment/selection of patients	People were recruited via local radio advertisements. Previous effectiveness of pharmaceutical interventions was not specifically assessed.
Age, gender and ethnicity	Age - Mean (range): 65.2 (50-86). Gender (M:F): 36:44. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Ranged from 2 to 30 years
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: NSAIDs - Etoricoxib. Etoricoxib 60mg once per day. Duration 14 days. Concurrent medication/care: People were asked to refrain from taking rescue medication within 12 hours (paracetamol) or 24 hours (Tramadol) of testing. Indirectness: No indirectness (n=40) Intervention 2: Placebo. Placebo once a day. Duration 14 days. Concurrent medication/care: People were asked to refrain from taking rescue medication within 12 hours (paracetamol) or 24 hours (Tramadol) of testing. Indirectness: No indirectness

Funding	Study funded by industry (This study was funded by an Investigator Initiated Study Programme grant from Merck Inc.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ETORICOXIB versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: Walk task pain (VAS) at 14 days; Group 1: mean 1.32 (SD 2.13); n=40, Group 2: mean 3.17 (SD 2.81); n=40; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline etoricoxib: 4.6 (1.4). Baseline placebo: 4.6 (1.5). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 6, Reason: Out of 86 people, 6 did not complete the study. 2 before starting, 1 withdrew for personal reasons, 3 due to adverse events; Group 2 Number missing: 6, Reason: Out of 86 people, 6 did not complete the study. 2 before starting, 1 withdrew for personal reasons, 3 due to adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Muller-fassbender 1994[136]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=200)
Countries and setting	Conducted in Italy; Setting: Inpatients
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Painful, active osteoarthritis of the knee diagnosis made based on Lequesne's clinical and radiographic criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with painful, active osteoarthritis of the knee. People presented with mild clinical signs of inflammation at the knee joint(s). They were of both sexes and over 18 years of age, with symptoms for at least 3 months and a Lequesne's index of at least 7 points. The knee joint involvement was either uni- or bilateral and the people had to be able to walk. People were hospitalised in two clinics involved in rehabilitation programs, mainly consisting of physical therapy.
Exclusion criteria	Significant haematological disorder at laboratory screening; history of hepatic or renal impairment; peptic ulcer disease; hypersensitivity to NSAIDs; recent injury to the involved knee(s) region; intraarticular corticosteroids within the previous 2 months; regular use of NSAIDs during the previous 2 months (the occasional use of NSAIDs as rescue medication for pain during this period was accepted). All the exclusions in the Lequesne's criteria with regard to rheumatic diseases other than osteoarthritis were systematically considered.
Recruitment/selection of patients	Recruited from two rheumatology clinics
Age, gender and ethnicity	Age - Mean (SD): 54 (8.5). Gender (M:F): 104:95. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 4.8 (3.4) years
Indirectness of population	No indirectness
Interventions	(n=100) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulfate 250mg - two tablets three times a day (1500mg per day). Duration 4 weeks. Concurrent medication/care: Concomitant

	intake of NSAIDs, corticosteroids, analgesics and other drugs for osteoarthritis was not permitted, while the person's program for physical therapy (including exercise, cold or heat application, etc.), if any, was allowed and had to be registered. Other treatments for concomitant diseases were allowed, provided that they were recorded.. Indirectness: No indirectness (n=100) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 200mg two tablets three times a day (1200mg/day). Duration 4 weeks. Concurrent medication/care: Concomitant intake of NSAIDs, corticosteroids, analgesics and other drugs for osteoarthritis was not permitted, while the person's program for physical therapy (including exercise, cold or heat application, etc.), if any, was allowed and had to be registered. Other treatments for concomitant diseases were allowed, provided that they were recorded.. Indirectness: No indirectness
Funding	Principal author funded by industry (Lucio C. Rovati and Ivo Setnikar were employees of the Department of Clinical Pharmacology, Rotta Research Laboratorium, Monza (MI), Italy)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus IBUPROFEN Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal complaints (including heartburn, gastric pain/discomfort, dyspepsia, nausea, early satiety, flatulence and diarrhoea, that were transient and predominantly of mild to moderate severity) at 4 weeks; Group 1: 5/100, Group 2: 29/99; Comments: No definition of individual types of adverse events inside of the broader defintion Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, localisation of osteoarthritis and symptom duration; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 was excluded after randomisation for reasons unrelated to the study and did not start the treatment - Actual outcome for Knee: Melaena at 4 weeks; Group 1: 0/100, Group 2: 1/99; Comments: 1 person complained of a couple of episodes of malaena, but gastroduodenoscopic investigation showed no evidence of bleeding Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, localisation of osteoarthritis and symptom duration; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 was excluded after randomisation for reasons unrelated to the study and did not start the treatment	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months;

	Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Munera 2010[137]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=315)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 5 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiologic evidence of osteoarthritis of the hip or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged 18 years and older with radiologic evidence of osteoarthritis of the hip or knee who had received opioid therapy in the previous year for osteoarthritis pain or whose osteoarthritis pain was inadequately controlled with NSAIDs
Exclusion criteria	People receiving more than 12 doses of short-acting oral opioids per day (or more than 90mg morphine equivalents per day); pregnant; allergic to opioids; allergic or otherwise intolerant to ibuprofen or skin adhesives; concomitant therapy with all analgesics (apart from aspirin as an antithrombotic at stable dosages ≤325mg/day), muscle relaxants or natural products for osteoarthritis pain, heat-based therapies and other investigational or unapproved drugs/devices
Recruitment/selection of patients	Before the study started there was a 1 week run-in period where previous medication was discontinued and people were maintained on 1600mg/day ibuprofen (400mg four times a day). If they had an average pain intensity of 7 or greater then they were permitted to enter the study.
Age, gender and ethnicity	Age - Mean (SD): 61.0 (12.7). Gender (M:F): 103:212. Ethnicity: White = 268, Black = 28, Hispanic = 16, Other = 3
Further population details	1. Age: Mixed (Includes people from 18-34 up to >80 years). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=152) Intervention 1: Opioids (topical - systemic) - Opioids (topical systemic). Buprenorphine transdermal system 5 microgram/hour uptitrated to 10-20

	microgram/hour as needed over a period of 21 days and then maintained for 7 days. Duration 4 weeks. Concurrent medication/care: Before the study started there was a 1 week run-in period where previous medication was discontinued and people were maintained on 1600mg/day ibuprofen (400mg four times a day). If they had an average pain intensity of 7 or greater then they were permitted to enter the study. On entry to the study the ibuprofen was discontinued. No rescue medication was allowed during this period.. Indirectness: No indirectness (n=163) Intervention 2: Placebo. Matching placebo patch. Duration 4 weeks. Concurrent medication/care: Before the study started there was a 1 week run-in period where previous medication was discontinued and people were maintained on 1600mg/day ibuprofen (400mg four times a day). If they had an average pain intensity of 7 or greater then they were permitted to enter the study. . Indirectness: No indirectness
Funding	Study funded by industry (This clinical research was funded by Purdue Pharma L.P.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OPIOIDS (TOPICAL SYSTEMIC) versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Other: Change from baseline in average pain intensity (numeric rating scale) at 4 weeks; Group 1: mean -1.84 (SD 2.69); n=149, Group 2: mean -1.4 (SD 2.67); n=162; Numeric rating scale 0-10 Top=High is poor outcome; Comments: Reports change scores and standard error. Reported buprenorphine: -1.84 (0.22). Reported placebo: -1.40 (0.21). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, predominant pain site, baseline medication use. Does not report baseline values for pain.; Group 1 Number missing: 83, Reason: 43 withdrew due to ineffective treatment, 36 due to adverse events, 2 due to protocol violations, 2 due to loss to follow up, 5 due to other; Group 2 Number missing: 77, Reason: 57 withdrew due to ineffective treatment, 18 due to adverse events, 4 due to protocol violation, 2 due to loss to follow up, 1 due to other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Niethard 2005[139]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=238)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinically diagnosed, symptomatic, unilateral osteoarthritis of the knee for at least 6 months confirmed at screening by radiographic observation of osteophytes and at least one of joint space narrowing, sclerosis, or subchondral cysts
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and females who were 45 years of age or older with clinically diagnosed, symptomatic, unilateral osteoarthritis of the knee for at least 6 months
Exclusion criteria	Clinically significant abnormalities in blood chemistry, haematology or urinalysis; radiographic evidence of severe OA (almost complete loss of joint space, large cysts, severe osteophytic changes, severe malformation of the joint); secondary osteoarthritis; history of rheumatoid arthritis or of any other chronic inflammatory disease such as colitis; history of fibromyalgia; current GI bleeding or history of bleeding over the last 3 years; significant injury to the target joint within 6 months prior to screening; major knee surgery of the target joint within 1 year of screening
Recruitment/selection of patients	After screening there was a washout phase for previous medication (lasting for at least 5 half-lives) where at the end people had to show that their knee was of at least moderate pain (more than or equal to 50mm on the visual analogue scale)
Age, gender and ethnicity	Age - Mean (SD): 66 (9). Gender (M:F): 57:56. Ethnicity: All participants were Caucasian
Further population details	1. Age: Mixed (Based on SD). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 6 months..
Indirectness of population	No indirectness

Interventions	(n=117) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Diclofenac diethylamine gel 1.16%, 4 grams applied four times a day - rubbed into the knee for no more than 1 minute until the gel vanished, paying specific attention to the medial area. Duration 3 weeks. Concurrent medication/care: People were permitted to use up to 4 tablets of rescue medication (paracetamol 500mg) per day for all pains they experienced regardless of origin. Indirectness: No indirectness (n=121) Intervention 2: Placebo. Placebo gel, 4 grams applied four times a day - rubbed into the knee for no more than 1 minute until the gel vanished, paying specific attention to the medial area. Duration 3 weeks. Concurrent medication/care: People were permitted to use up to 4 tablets of rescue medication (paracetamol 500mg) per day for all pains they experienced regardless of origin. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Novartis Consumer Health)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 3 weeks; Group 1: mean -22 (SD 21); n=117, Group 2: mean -14 (SD 23); n=120; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 48 (16). Baseline placebo: 47 (16).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, previous treatment, position of pain, presence of joint effusion, X-ray findings and baseline values for outcomes; Group 1 Number missing: 15, Reason: 2 adverse events, 1 lack of efficacy, 10 protocol violations, 1 withdrew consent, 1 administrative; Group 2 Number missing: 23, Reason: 2 lack of efficacy, 16 protocol violations, 3 withdrew consent, 2 administrative

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 3 weeks; Group 1: mean -23 (SD 21); n=117, Group 2: mean -16 (SD 22); n=120; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 53 (15). Baseline placebo: 51 (15).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, previous treatment, position of pain, presence of joint effusion, X-ray findings and baseline values for outcomes; Group 1 Number missing: 15, Reason: 2 adverse events, 1 lack of efficacy, 10 protocol violations, 1 withdrew consent, 1 administrative; Group 2 Number missing: 23, Reason: 2 lack of efficacy, 16 protocol violations, 3 withdrew consent, 2 administrative

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events (dry mouth and nausea) at 3 weeks; Group 1: 0/117, Group 2: 2/121

Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, previous treatment, position of pain, presence of joint effusion, X-ray findings and baseline values for outcomes; Group 1 Number missing: 15, Reason: 2 adverse events, 1 lack of efficacy, 10 protocol violations, 1 withdrew consent, 1 administrative; Group 2 Number missing: 23, Reason: 2 lack of efficacy, 16 protocol violations, 3 withdrew consent, 2 administrative Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Brain tumour at 3 weeks; Group 1: 0/117, Group 2: 1/121 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, gender, ethnicity, previous treatment, position of pain, presence of joint effusion, X-ray findings and baseline values for outcomes; Group 1 Number missing: 15, Reason: 2 adverse events, 1 lack of efficacy, 10 protocol violations, 1 withdrew consent, 1 administrative; Group 2 Number missing: 23, Reason: 2 lack of efficacy, 16 protocol violations, 3 withdrew consent, 2 administrative	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Noack 1994[140]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=252)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee (gonarthritis) that was uni or bilateral with pain and limitation of motion but without overt laboratory or clinical signs of inflammation at the affected joint. Osteoarthritis of the knee was defined according to the clinical and radiological criteria of Lequesne.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Ambulatory people of both sexes with osteoarthritis of the knee who were over 18 years of age with no overt laboratory or clinical signs of inflammation at the affected joints.
Exclusion criteria	A radiological stage below I or above III; intra-articular corticosteroids within the previous 2 months; NSAID use within the previous 2 weeks; recent trauma or lesions of the involved knee(s); significant haematological disorder, hepatic or renal abnormality at laboratory screening; extreme under- or overweight (Broca index <75 and >150); all exclusions present in Lequesne's criteria were systematically considered, with particular regard to inflammatory rheumatic diseases, metabolic arthropathies, and, in any case, rheumatic diseases other than osteoarthritis.
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 55 (14.5). Gender (M:F): 100:152. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range (17-85)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Radiological stage (Jager and Wirth) I-III, clinical stage (Weseloh and Liebig) 1-4 Duration of symptoms: Between <6 months and >10 years. Median 2-10 years..
Indirectness of population	No indirectness
Interventions	(n=126) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). 250mg glucosamine sulfate two tablets three times a day

	(1500mg/day). Duration 4 weeks. Concurrent medication/care: NSAIDs, analgesics, corticosteroids, or other treatments for osteoarthritis, including any physical therapy, were not allowed during the study. Other treatments for concomitant diseases were allowed, but had to be recorded.. Indirectness: No indirectness (n=126) Intervention 2: Placebo. Placebo two tablets three times a day. Duration 4 weeks. Concurrent medication/care: NSAIDs, analgesics, corticosteroids, or other treatments for osteoarthritis, including any physical therapy, were not allowed during the study. Other treatments for concomitant diseases were allowed, but had to be recorded.. Indirectness: No indirectness
Funding	Principal author funded by industry (Lucio C. Rovati and Ivo Setnikar were employees of the Department of Clinical Pharmacology, Rotta Research Laboratrium, Monza, Italy. Micheal Fischer and Klaus K. Forster were employees of GmBH.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disturbances (including epigastric pain, discomfort, nausea, vomiting, diarrhoea) at 4 weeks; Group 1: 5/126, Group 2: 6/126; Comments: No defintiion of the number of people in each category
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, localisation of osteoarthritis, radiological stage, clinical stage and duration of symptoms; Group 1 Number missing: 4; Reason: 1 discontinued due to adverse events, 3 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.; Group 2 Number missing: 4; Reason: 2 discontinued due to adverse events, 2 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Low blood pressure, recurrent heart failure at 4 weeks; Group 1: 0/126, Group 2: 2/126; Comments: Placebo: low blood pressure = 1, recurrent heart failure = 1
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, localisation of osteoarthritis, radiological stage, clinical stage and duration of symptoms; Group 1 Number missing: 4; Reason: 1 discontinued due to adverse events, 3 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.; Group 2 Number missing: 4; Reason: 2 discontinued due to adverse events, 2 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.
Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headaches at 4 weeks; Group 1: 0/126, Group 2: 2/126
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, localisation of osteoarthritis,

radiological stage, clinical stage and duration of symptoms; Group 1 Number missing: 4; Reason: 1 discontinued due to adverse events, 3 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.; Group 2 Number missing: 4; Reason: 2 discontinued due to adverse events, 2 did not return to follow up. An additional 3 people between both groups were excluded due to protocol violation.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Nowlan 2003[141]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of osteoarthritis in any joint (with the exception of osteoarthritis solely in the spine) by their primary care physicians using either x-ray examinations or clinical criteria
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People 50 years or older; currently being treated with nonsteroidal anti-inflammatory drugs or paracetamol for osteoarthritis pain; previously diagnosed with osteoarthritis in any joint (with the exception of osteoarthritis soley in the spine) by their primary care physicians using either x-ray examinations or clinical criteria
Exclusion criteria	Previous gastrointestinal bleeding or known gastrointestinal ulcer; any other rheumatic disease; intra-articular injection; known allergy to non-steroidal anti-inflammatory drugs or glucosamine sulfate; current participation in another study evaluating osteoarthritis treatments; regularly taking glucosamine sulfate currently or for a period of 6 weeks or longer within the past 3 months.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 65.2 (9.0). Gender (M:F): 10:30. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Mixed 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulfate 500mg three times a day. Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg to be used as necessary as an adjuvant medicine for relief of arthritis pain. Indirectness: No indirectness (n=20) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 400mg three times daily.

	Duration 12 weeks. Concurrent medication/care: Paracetamol 500mg to be used as necessary as an adjuvant medicine for relief of arthritis pain. Indirectness: No indirectness
Funding	Study funded by industry ("We thank the College of Family Physicians of Canada for the first author's scholarship; Physician's Service Incorporated Foundation for financial assistance; Swiss Herbal Remedies, McNeil Consumer Products, and AMCO Pharmaceuticals for in-kind contributions; Anne Grindrod for statistical analysis and data entry; the staff of the Byron and Victoria Family Medicine Centers for their assistance; and Drs Moira Stewart and Beth Henning for suggestions of revising this article". Drugs supplied by industry.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus IBUPROFEN Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Other: Dyspepsia at 12 weeks; Group 1: 1/20, Group 2: 3/20 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, joints involved, previously used class of analgesic, baseline pain scores; Group 1 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief; Group 2 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Palpitations at 12 weeks; Group 1: 1/20, Group 2: 0/20 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, joints involved, previously used class of analgesic, baseline pain scores; Group 1 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief; Group 2 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Vertigo, tinnitus at 12 weeks; Group 1: 0/20, Group 2: 2/20; Comments: Ibuprofen: Vertigo = 1, tinnitus = 1 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, joints involved, previously used class of analgesic, baseline pain scores; Group 1 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief; Group 2 Number missing: 1, Reason: 1 withdrew early due to inadequate pain relief	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months;

	Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	PACES trial: Pincus 2004[148]
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	2 (pooled analysis of 2 identical studies) (n=1080)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention time: 6 weeks for each intervention (14 weeks including washout periods in total)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee or hip with pain (40-90mm on a visual analogue pain scale) and Kellgren-Lawrence grade 2-4 changes
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 45 or greater, radiographic Kellgren-Lawrence grade 2-4, a score of 40-90mm on a visual analogue pain scale, and designation by the treating physician that the patient was a candidate for long term treatment with a cyclo-oxygenase-2 specific inhibitor drug or an analgesic drug
Exclusion criteria	Significant medical comorbidities; rheumatoid arthritis or other inflammatory arthritis; acute joint trauma; chronic pain syndrome; expected need for surgery during the course of the study; oral or parenteral corticosteroids within 2 months; intra-articular injections of hyaluronic acid within 9 months; women of childbearing potential who did not use contraception; pregnant or lactating women
Recruitment/selection of patients	Two separate trials
Age, gender and ethnicity	Age - Mean (SD): 63.45 (10.03). Gender (M:F): 367:713. Ethnicity: 86.1% were Caucasian, no information about the remainder of participants
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip (majority knee)).
Extra comments	Severity: Kellgren Lawrence mean radiographic grade 2-3 Duration of symptoms (mean [SD]): 9.36 (8.99) years
Indirectness of population	No indirectness
Interventions	(n=723) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg/day taken in the morning . Duration 6 weeks. Concurrent medication/care: Propoxyphene 65mg up to four times a day was given as rescue treatment; codeine 60mg or tramadol 100mg up

	to four times per day were provided as alternatives to fewer than 5% of people if propoxyphene was poorly tolerated or ineffective. Indirectness: No indirectness (n=631) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg four times a day. Duration 6 weeks. Concurrent medication/care: Propoxyphene 65mg up to four times a day was given as rescue treatment; codeine 60mg or tramadol 100mg up to four times per day were provided as alternatives to fewer than 5% of people if propoxyphene was poorly tolerated or ineffective. Indirectness: No indirectness (n=562) Intervention 3: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: Propoxyphene 65mg up to four times a day was given as rescue treatment; codeine 60mg or tramadol 100mg up to four times per day were provided as alternatives to fewer than 5% of people if propoxyphene was poorly tolerated or ineffective. Indirectness: No indirectness
Funding	Study funded by industry (Sponsored by Pfizer Corporation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: MDHAQ visual analogue pain scale at 6 weeks; Group 1: mean -8.09 (SD 35.3); n=709, Group 2: mean -5.47 (SD 29.21); n=603; MDHAQ (Multidimensional Health Assessment Questionnaire) visual analogue pain scale 0-100 Top=High is poor outcome; Comments: Values reported for each treatment arm (ex. celecoxib-paracetamol, celecoxib-placebo). Values combined for the result. Reports mean and standard error. Converted to standard deviation. PACES-a: Reported celecoxib-paracetamol: Period 1 = -3.47 (2.66). Period 2 = -6.49 (2.17). Reported celecoxib-placebo: Period 1 = -9.25 (2.63). Period 2 = -7.66 (2.47). Reported paracetamol-placebo: Period 1 = -5.78 (2.68). Period 2 = -1.17 (2.28). PACES-b: Reported celecoxib-paracetamol: Period 1 = -7.69 (2.54). Period 2 = -4.11 (2.13). Reported celecoxib-placebo: Period 1 = -13.84 (2.54). Period 2 = -9.98 (2.43). Reported paracetamol-placebo: Period 1 = -6.15 (2.55). Period 2 = -5.87 (2.54).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 28, Reason: Reasons not given

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Any gastrointestinal event at 6 weeks; Group 1: 77/723, Group 2: 58/631; Comments: PACES-a: Celecoxib (n=350): any gastrointestinal event = 42; diarrhoea = 8; dyspepsia = 10; nausea = 7; flatulence = 8. Paracetamol (n=300): any gastrointestinal event = 28; diarrhoea = 14; dyspepsia = 7; nausea = 7; flatulence = 4. PACES-b: Celecoxib (n=373): any gastrointestinal event = 35; diarrhoea = 6; dyspepsia = 6; nausea = 8; flatulence = 3. Paracetamol (n=331): any gastrointestinal event = 30; diarrhoea = 11; dyspepsia = 6; nausea = 4; flatulence = 4.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 28, Reason: Reasons not given

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Headache at 6 weeks; Group 1: 11/723, Group 2: 20/631; Comments: PACES-a: Celecoxib: 4, paracetamol: 11. PACES-b: Celecoxib: 7, paracetamol: 9
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 28, Reason: Reasons not given

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: MDHAQ visual analogue pain scale at 6 weeks; Group 1: mean -8.09 (SD 35.3); n=709, Group 2: mean -6.79 (SD 28.63); n=519; MDHAQ visual analogue scale 0-100 Top=High is poor outcome; Comments: Values reported for each treatment arm (ex. celecoxib-paracetamol, celecoxib-placebo). Values combined for the result. Reports mean and standard error. Converted to standard deviation. PACES-a: Reported celecoxib-paracetamol: Period 1 = -3.47 (2.66). Period 2 = -6.49 (2.17). Reported celecoxib-placebo: Period 1 = -9.25 (2.63). Period 2 = -7.66 (2.47). Reported paracetamol-placebo: Period 1 = -5.78 (2.68). Period 2 = -1.17 (2.28). PACES-b: Reported celecoxib-paracetamol: Period 1 = -7.69 (2.54). Period 2 = -4.11 (2.13). Reported celecoxib-placebo: Period 1 = -13.84 (2.54). Period 2 = -9.98 (2.43). Reported paracetamol-placebo: Period 1 = -6.15 (2.55). Period 2 = -5.87 (2.54).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Any gastrointestinal event at 6 weeks; Group 1: 77/723, Group 2: 43/562; Comments: PACES-a: Celecoxib (n=350): any gastrointestinal event = 42; diarrhoea = 8; dyspepsia = 10; nausea = 7; flatulence = 8. Placebo (n=289): any gastrointestinal event = 26; diarrhoea = 4; dyspepsia = 3; nausea = 5; flatulence = 1. PACES-b: Celecoxib (n=373): any gastrointestinal event = 35; diarrhoea = 6; dyspepsia = 6; nausea = 8; flatulence = 3. Placebo (n=273): any gastrointestinal event = 17; diarrhoea = 4; dyspepsia = 2; nausea = 3; flatulence = 1.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Headache at 6 weeks; Group 1: 11/723, Group 2: 12/562; Comments: PACES-a: Celecoxib: 4, placebo: 5. PACES-b: Celecoxib: 7,

placebo: 7
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 14, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: MDHAQ visual analogue pain scale at 6 weeks; Group 1: mean -5.47 (SD 29.21); n=603, Group 2: mean -6.79 (SD 28.63); n=519; MDHAQ visual analogue pain scale 0-100 Top=High is poor outcome; Comments: Values reported for each treatment arm (ex. celecoxib-paracetamol, celecoxib-placebo). Values combined for the result. Reports mean and standard error. Converted to standard deviation. PACES-a: Reported celecoxib-paracetamol: Period 1 = -3.47 (2.66). Period 2 = -6.49 (2.17). Reported celecoxib-placebo: Period 1 = -9.25 (2.63). Period 2 = -7.66 (2.47). Reported paracetamol-placebo: Period 1 = -5.78 (2.68). Period 2 = -1.17 (2.28). PACES-b: Reported celecoxib-paracetamol: Period 1 = -7.69 (2.54). Period 2 = -4.11 (2.13). Reported celecoxib-placebo: Period 1 = -13.84 (2.54). Period 2 = -9.98 (2.43). Reported paracetamol-placebo: Period 1 = -6.15 (2.55). Period 2 = -5.87 (2.54).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 28, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Any gastrointestinal event at 6 weeks; Group 1: 58/631, Group 2: 43/562; Comments: PACES-a: Paracetamol (n=300): any gastrointestinal event = 28; diarrhoea = 14; dyspepsia = 7; nausea = 7; flatulence = 4. Placebo (n=289): any gastrointestinal event = 26; diarrhoea = 4; dyspepsia = 3; nausea = 5; flatulence = 1. PACES-b: Paracetamol (n=331): any gastrointestinal event = 30; diarrhoea = 11; dyspepsia = 6; nausea = 4; flatulence = 4. Placebo (n=273): any gastrointestinal event = 17; diarrhoea = 4; dyspepsia = 2; nausea = 3; flatulence = 1.
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI distress; Group 1 Number missing: 28, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Headache at 6 weeks; Group 1: 20/631, Group 2: 12/562; Comments: PACES-a: Paracetamol: 11, placebo: 5. PACES-b: Paracetamol: 9, placebo: 7
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, percentage who were Caucasian, education, duration of disease, radiographic grade, knee as index joint, taking NSAIDs prior to trial, WOMAC target joint, MDHAQ VAS and MDHAQ GI

distress; Group 1 Number missing: 28, Reason: Reasons not given; Group 2 Number missing: 43, Reason: Reasons not given	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months

Study	Paul 2009[142]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=423)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Uncomplicated osteoarthritis of the knee joint with symptoms for at least 6 months having fulfilled the X-ray criteria for Kellgren's classes II or III and those who met American Rheumatologic Association functional class I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of either sex, 40-64 years of age, with uncomplicated osteoarthritis of the knee joint with symptoms for at least 6 months having fulfilled the X-ray criteria for Kellgren's classes II or III and those who met American Rheumatologic Association functional class I-III
Exclusion criteria	People with significant renal impairment (creatinine clearance <30 mL/min); class III/IV angina; uncontrolled congestive heart failure; uncontrolled hypertension (BP >160/100); clinically significant physical and mental abnormalities or abnormal laboratory examinations; hepatic disease; bronchial asthma; a history of active active gastrointestinal bleeding; neoplasia; acute meniscus injury, arthroscopy in the study joint within the last 6 months; obesity (BMI no less than 40); those allergic to conventional NSAIDs; people requiring systemic steroids, warfarin, lithium, low dose aspirin, anti-ulcer drugs or intra-articular steroids within the last 2 months of entering the study
Recruitment/selection of patients	No additional criteria
Age, gender and ethnicity	Age - Mean (range): 53.5 (40-64). Gender (M:F): 189:234. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren and Lawrence grade 2-3. American Rheumatologic Association functional class 1-3. Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness

Interventions	(n=282) Intervention 1: NSAIDs - Aceclofenac. Aceclofenac (100mg) or nabumetone (750mg) twice daily. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The aceclofenac and nabumetone groups were merged due to class effect as agreed in the protocol (n=141) Intervention 2: Placebo. Matching placebo twice daily. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACECLOFENAC OR NABUMETONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 4 weeks; Group 1: mean -4.02 (SD 1.79); n=226, Group 2: mean -0.79 (SD 1.78); n=89; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Reports mean change scores and standard error. Converted into standard deviation. Reported aceclofenac: 3.79 (0.17). Reported nabumetone: 4.23 (0.13). Reported placebo: 0.79 (0.15). Calculated SD aceclofenac: 2.01. Calculated SD nabumetone: 1.54. Calculated SD placebo: 1.78.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values for sex and age, but does not report individual group values or baseline values for outcomes; Group 1 Number missing: 56, Reason: 108 completed the full course of treatment with aceclofenac, 118 completed the full course of treatment with nabumetone. Reason unclear (some due to adverse events).; Group 2 Number missing: 52, Reason: 89 completed the full course of treatment with placebo. Reason unclear (some due to adverse events).

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal side effects at 4 weeks; Group 1: 18/282, Group 2: 3/141; Comments: Aceclofenac: 11. Nabumetone: 7. Placebo: 3.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Only reports adverse events leading to discontinuation from the study; Baseline details: Reports overall values for sex and age, but does not report individual group values or baseline values for outcomes; Group 1 Number missing: 56, Reason: 108 completed the full course of treatment with aceclofenac, 118 completed the full course of treatment with nabumetone. Reason unclear (some due to adverse events).; Group 2 Number missing: 52, Reason: 89 completed the full course of treatment with placebo. Reason unclear (some due to adverse events).

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-

	months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Pavelka 1998[144]
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Czech Republic, Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks for each intervention arm
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiologically confirmed osteoarthritis (Kellgren grade 2-4) of the hip and/or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females age over 18 years with pain associated with radiologically confirmed osteoarthritis (Kellgren grade 2-4) of the hip and/or knee necessitating analgesic treatment with the pain at least moderate on a one off 4-point verbal rating scale (0 = none, 3 = severe).
Exclusion criteria	Presence of gout; inflammatory arthropathy or any other painful conditions likely to affect the results; a history of epilepsy; clinically significant abnormal laboratory results; impairment of any organ system likely to prohibit the use of tramadol or diclofenac; hypersensitivity to either drug; pregnant or lactating women; people who were considered to be suicidal; people with a history of substance abuse; people requiring concurrent physiotherapy or treatment with monoamine oxidase inhibitors, anxiolytic or sedative drugs (except for use as night sedation); people who received any unpermitted medication less than 30 days prior to the study; people who received intra-articular corticosteroid injections in the 3 months prior to study recruitment
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Range: 44 to 85 years. Gender (M:F): 8:52. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on range). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (19) and knee (41)).
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms (median [IQR]): group 1 = 5.25 (3.00 to 10.00) years, group 2 = 8.00 (3.00 to 10.00) years
Indirectness of population	No indirectness

Interventions	(n=60) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol 50mg capsules - 1-2 capsules up to 3 times a day (as required for pain relief). Duration 4 weeks. Concurrent medication/care: During the washout period people were only allowed to take paracetamol. No other analgesic drugs were permitted during the study.. Indirectness: No indirectness (n=60) Intervention 2: NSAIDs - Diclofenac. Diclofenac 25mg tablets - 1-2 tablets up to 3 times a day (as required for pain relief). Duration 4 weeks. Concurrent medication/care: During the washout period people were only allowed to take paracetamol. No other analgesic drugs were permitted during the study.. Indirectness: No indirectness
Funding	Study funded by industry (This study was sponsored by Grünenthal GmbH, Aachen, Germany)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus DICLOFENAC Protocol outcome 1: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 4 weeks; Group 1: 4/60, Group 2: 0/60 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports WOMAC score at baseline, duration of arthritis, type of osteoarthritis and age.; Group 1 Number missing: 0, Reason: All people who withdrew from the study withdrew for adverse events accounted for in the outcome; Group 2 Number missing: 0, Reason: All people who withdrew from the study withdrew for adverse events accounted for in the outcome	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study (subsidiary papers)	Pavelka 2002[143] (Bruyere 2004[38])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=202)
Countries and setting	Conducted in Czech Republic; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis of the medial femorotibial compartment based on the clinical and radiological criteria of the American College of Rheumatology
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of both sexes aged 45 to 70 years with primary knee osteoarthritis. Minimum symptom severity was ensured by using a Lequesne algo-functional index score of at least 4 points.
Exclusion criteria	Lequesne score greater than 12; history of clinically significant articular and rheumatic diseases, or that may cause secondary arthritis, including a history of trauma or lesions of the knee joint and severe articular inflammation as confirmed by physical examination (eg, a finding of severe joint effusion); evidence of rapidly progressive osteoarthritis obtained before the trial; overweight, defined as a body mass index greater than 27; clinically significant alterations in haematologic variables and renal, hepatic, and metabolic functions in the opinion of the investigator (including a history of clinically evidence diabetes mellitus); systemic or intra-articular corticosteroid therapy in the previous 3 months
Recruitment/selection of patients	Single centre trial
Age, gender and ethnicity	Age - Mean (SD): 62.4 (7.2). Gender (M:F): 45:157. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren and Lawrence grade 2-3 Duration of symptoms: 10.6 (7.5) years
Indirectness of population	No indirectness
Interventions	(n=101) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulfate formulation (packets of powder for oral solution) with a net content equivalent to 1500mg of glucosamine sulfate, one per

	day. Duration 3 years. Concurrent medication/care: Paracetamol in 500mg tablets was provided for rescue medication as needed, and its use was recorded in a patient daily diary. No other pharmacologic treatments for osteoarthritis or other formulations containing analgesics were allowed. Among physical therapies, only hydroptherapy, exercise and ultrasound, alone or in combination were allowed if the person was following a stable regimen. Indirectness: No indirectness (n=101) Intervention 2: Placebo. One placebo sachet of powder for solution per day. Duration 3 years. Concurrent medication/care: Paracetamol in 500mg tablets was provided for rescue medication as needed, and its use was recorded in a patient daily diary. No other pharmacologic treatments for osteoarthritis or other formulations containing analgesics were allowed. Among physical therapies, only hydroptherapy, exercise and ultrasound, alone or in combination were allowed if the person was following a stable regimen. Indirectness: No indirectness
Funding	Principal author funded by industry (Drs Giacovelli and Rovati were employees at the Department of Clinical Pharmacology, Rotta Research Laboratorium, Monza, Italy)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 3 years; Group 1: mean -2 (SD 2.3); n=101, Group 2: mean -1.3 (SD 6.6); n=101; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Calculated SD from this. Reported glucosamine: -2.0 (-2.4 to -1.5). Reported placebo: -1.3 (-1.7 to 0.88). Baseline glucosamine: 6.61 (3.45). Baseline placebo: 6.33 (3.13).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, Kellgren and Lawrence grade, joint space width, Lequesne index score and WOMAC index score; Group 1 Number missing: 35, Reason: 35 withdrew. 8 due to adverse events, 16 free choice, 8 intervention ineffective, 3 lost to follow up.; Group 2 Number missing: 46, Reason: 46 withdrew. 10 due to adverse events, 26 free choice, 5 intervention ineffective, 5 lost to follow up.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC function subscale at 3 years; Group 1: mean -5.8 (SD 6.9); n=101, Group 2: mean -3.7 (SD 6.2); n=101; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Calculated SD from this. Reported glucosamine: -5.8 (-7.1 to -4.4). Reported placebo: -3.7 (-4.9 to -2.5). Baseline glucosamine: 21.84 (10.67). Baseline placebo: 22.00 (11.03).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, Kellgren and Lawrence grade, joint space width, Lequesne index score and WOMAC index score; Group 1 Number missing: 35, Reason: 35 withdrew. 8 due to adverse events, 16 free choice, 8 intervention ineffective, 3 lost to follow up.; Group 2 Number missing: 46, Reason: 46 withdrew. 10 due to adverse events, 26 free

choice, 5 intervention ineffective, 5 lost to follow up. Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiovascular adverse events at 3 years; Group 1: 23/101, Group 2: 20/101; Comments: Reported glucosamine: 23%. Reported placebo: 20%. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, Kellgren and Lawrence grade, joint space width, Lequesne index score and WOMAC index score; Group 1 Number missing: 35, Reason: 35 withdrew. 8 due to adverse events, 16 free choice, 8 intervention ineffective, 3 lost to follow up.; Group 2 Number missing: 46, Reason: 46 withdrew. 10 due to adverse events, 26 free choice, 5 intervention ineffective, 5 lost to follow up.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Peloso 2000[145]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=103)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis requiring symptoms, including pain, stiffness, and disability, that required the use of paracetamol, antiinflammatory agents or opioid analgesics for the previous 3 months of longer. Radiographic confirmation of the minimum of grade II osteoarthritis severity of the hip or knee joint.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with primary osteoarthritis who were male or nonpregnant females and over the age of 35 years. People also needed to have grade 2 osteoarthritis severity of a knee or hip joint, as defined by the standard atlas of radiographs. Grade 2 changes require the presence of an osteophyte and joint space narrowing. People with more advanced radiographic grades were also eligible provided that imminent joint replacement surgery would not prevent study completion.
Exclusion criteria	Known allergy to codeine, other opioids or paracetamol; history of previous opioid abuse, characterised by compulsive drug use, an intense desire for the drug, and an overwhelming craving for its continued availability, or if manipulation of a previous physician or the medical system for the purposes of obtaining additional drug was suspected; secondary causes of ostearthritis; if they had received systemic or intraarticular corticosteroids in the past 2 months or intraarticular viscosupplementation in the past 6 months; people with grade 4 severity and awaiting replacement surgery
Recruitment/selection of patients	Four Canadian sites. People were required to discontinue their prestudy analgesics, and experience a flare in their hip and/or knee during a 2-7 day washout period. A flare was defined as an increase in pain to a minimum report of moderate pain on a 5 point Likert scale (none, mild, moderate, severe and excruciating pain).
Age, gender and ethnicity	Age - Mean (SD): 62.2 (10.5). Gender (M:F): 39:64. Ethnicity: Not stated

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (49) or knee (94) or both).
Extra comments	Severity: At least radiographic grade II Duration of symptoms (mean [SD]): 10.3 (7.5) years.
Indirectness of population	No indirectness
Interventions	(n=51) Intervention 1: Weak opioids (oral) - Codeine. Controlled release codeine administered 12 hourly. Treatment was initiated at a dose of 100mg of controlled release codeine per day and the dose was escalated weekly, provided there was ongoing pain and a lack of limiting side effects, up to a maximum of 400mg per day.. Duration 4 weeks. Concurrent medication/care: Use of additional anti-inflammatory or analgesic medication, other than paracetamol 650mg up to 3 times daily for control of pain not managed by controlled release codeine or placebo, was not permitted.. Indirectness: No indirectness (n=52) Intervention 2: Placebo. Matching placebo twice daily. Duration 4 weeks. Concurrent medication/care: Use of additional anti-inflammatory or analgesic medication, other than paracetamol 650mg up to 3 times daily for control of pain not managed by controlled release codeine or placebo, was not permitted.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CODEINE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC VAS pain at 4 weeks; Group 1: mean 118 (SD 106.3); n=31, Group 2: mean 31.1 (SD 92); n=35; WOMAC pain visual analogue scale 0-500 Top=High is poor outcome; Comments: Baseline codeine: 263.5 (99.7). Baseline placebo: 252.4 (120.8).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, weight, height, type of osteoarthritis, and baseline values for outcomes; Group 1 Number missing: 20, Reason: 20 did not complete. 15 due to adverse events, 1 due to unrelated illness, 1 due to patient noncompliance, 1 due to patient withdrawal, 1 due to other reasons.; Group 2 Number missing: 17, Reason: 17 did not complete. 4 due to adverse events, 5 due to inadequate pain control, 1 due to patient noncompliance, 1 due to patient withdrawal, 1 due to protocol violation, 5 due to other reasons

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function at 4 weeks; Group 1: mean 444.2 (SD 400.8); n=31, Group 2: mean 143.5 (SD 284.7); n=35; WOMAC

physical function 0-1700 Top=High is poor outcome; Comments: Baseline codeine: 900.5 (357.3). Baseline placebo: 844.9 (405.3). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, duration of osteoarthritis, weight, height, type of osteoarthritis, and baseline values for outcomes; Group 1 Number missing: 20, Reason: 20 did not complete. 15 due to adverse events, 1 due to unrelated illness, 1 due to patient noncompliance, 1 due to patient withdrawal, 1 due to other reasons.; Group 2 Number missing: 17, Reason: 17 did not complete. 4 due to adverse events, 5 due to inadequate pain control, 1 due to patient noncompliance, 1 due to patient withdrawal, 1 due to protocol violation, 5 due to other reasons	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	**Persson 2021[147]**
Study type	RCT (Patient randomised; Crossover: 4 weeks)
Number of studies (number of participants)	1 (n=22)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 week intervention trials with each drug for three cycles (total: 24 weeks on treatment + 20 weeks maximum for washout periods)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Chronic knee pain and radiographic knee osteoarthritis (i.e. definite narrowing and definite osteophyte in the tibiofemoral and/or patellofemoral compartments as per Nottingham line drawing atlas scoring.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged 40 years and over with chronic knee pain and radiographic knee osteoarthritis; people scoring between 4 and 8 on the 0-10 numeric rating scale.
Exclusion criteria	Inability to give informed consent; terminal or untreated major mental illness; pregnancy or breastfeeding; daily use of oral NSAID in the last 2 weeks; prior regular use of ibuprofen gel or capsaicin cream on the affected knee(s); hypersensitivity or

	allergy to the interventions or other ingredients in the preparations; total joint replacement of the affected joint; current treatment for stomach or duodenal ulcers; renal failure; current treatment with anticoagulants.
Recruitment/selection of patients	People were recruited from the Nottingham Knee Pain and Health in the Community cohort study from the East Midlands region of the UK
Age, gender and ethnicity	Age - Mean (SD): 67.0 (9.3). Gender (M:F): 10:12. Ethnicity: Not stated/unclear
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Low comorbidity score (9 had increased scores for HADS anxiety subscale, 4 had increased scores for HADS depression subscale, 2 met the criteria for fibromyalgia). 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Total NLDA score (median [IQR]) = 13 (9 to 18) Duration of symptoms: Not stated/unclear. NCT03146689
Indirectness of population	No indirectness
Interventions	(n=22) Intervention 1: NSAID gels (topical - local) - Ibuprofen gel. 5% w/w ibuprofen gel. Applied four times daily to the painful knee(s). The recommended dose was an extruded inch of ibuprofen gel.. Duration 12 weeks. Concurrent medication/care: People continued to use their regular medications, including oral analgesics, throughout the trial provided the frequency/dose had remained stable for 3 months. Non-permitted concomitant therapies were additional topical analgesics for the affected knee, regular oral NSAIDs, joint infection or surgery.. Indirectness: No indirectness Comments: 3 x 4 weeks trials with ibuprofen would take place (n=22) Intervention 2: Capsaicin cream (topical - local) - Capsaicin cream (topical). 0.025% w/w capsaicin cream applied four times daily to the painful knee(s). The recommended dose was a pea-sized amount of capsaicin cream.. Duration 12 weeks. Concurrent medication/care: People continued to use their regular medications, including oral analgesics, throughout the trial provided the frequency/dose had remained stable for 3 months. Non-permitted concomitant therapies were additional topical analgesics for the affected knee, regular oral NSAIDs, joint infection or surgery.. Indirectness: No indirectness Comments: 3 x 4 weeks trials with ibuprofen would take place
Funding	Other author(s) funded by industry (M.S.M.P declares a grant from the Nottingham University Charity for the conduct of the reported work; J.S. reports grants from Vs Arthritis and Nottingham University Charity during the conduct of the study; D.A.W. reports grants from Vs Arthritis and Nottingham University Charity during the conduct of the study; and other from AbbVie Ltd, Pfizer Ltd, Eli Lilly & Co Ltd, GlaxoSmithKline Medscape Education (New York), Love Productions (UK) outside the submitted work; M.D. reports grants from Vs Arthritis and from Nottingham University Hospital Charitable Trust during the conduct of the study; W.Z. reports grants from Vs Arthritis and Nottingham University Hospital Charity during the conduct of the study; and personal fees from Grunenthal and Regeneron Inc. outside the submitted work.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN GEL versus CAPSAICIN CREAM (TOPICAL)

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain reduction (NRS) at 12 weeks; Group 1: mean -1.2 (SD 0.3); n=22, Group 2: mean -1.6 (SD 0.4); n=22; NRS 0-10 Top=High

is poor outcome; Comments: Reported mean change score and 95% confidence interval. Reported ibuprofen: 1.2 (0.5-1.8). Reported capsaicin: 1.6 (0.9-2.4). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, BMI, comorbidities and osteoarthritis features; Group 1 Number missing: 5, Reason: 5 did not complete the trial (3 treatment related adverse events, 1 death (unrelated), 1 consent withdrawn).; Group 2 Number missing: 5, Reason: 5 did not complete the trial (3 treatment related adverse events, 1 death (unrelated), 1 consent withdrawn).	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Prior 2014[150]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=542)
Countries and setting	Conducted in USA; Setting: Private, ambulatory, primary care sites
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the knee had to have knee pain, radiographic osteophytes and at least one of: morning stiffness of less than 30 minutes duration, crepitus on motion, or were 40 or more years of age. People with osteoarthritis of the hip had to have hip pain, radiographic femoral and/or acetabular osteophytes, and radiographic joint space narrowing as established by the American College of Rheumatology for idiopathic osteoarthritis of the hip. Had to have a radiograph taken within 6 months showing Kellgren Lawrence grade 2-3 changes.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women 40 years of age or older who met the following clinical criteria: symptomatic idiopathic osteoarthritis of the hip or knee for a minimum of 6 months; osteoarthritis pain that required the use of nonsteroidal anti-inflammatory drugs, paracetamol, or another analgesic agent on a regular basis (three or more days per week); history of positive therapeutic benefit with paracetamol used for osteoarthritis pain (this was considered analogous to 'flare' design used in many trials of anti-inflammatory drugs); and history of moderate to severe osteoarthritis of the hip or knee when not taking osteoarthritis analgesic medication; Kellgren Lawrence grade 2-3 joint changes; ACE functional class I-III; could walk at least 100 feet without an assisting device; any knee joint effusion had to be aspirated and found noninflammatory; alkaline phosphatase ≤1.5 times the upper limit of normal; alanine aminotransferase, aspartate aminotransferase and blood urea nitrogen ≤1.5 times the upper limit of normal and serum creatinine ≤1.5mg/dL. Other routine haematologic, biochemical, and urinalysis laboratory parameters were within normal range, or if outside normal range, were judged not clinically significant. Women were post-menopausal or using an effective form of birth control for at least 3 months before study entry and during the study.
Exclusion criteria	History of surgery, including arthroscopy, or major trauma to the study joint in the past 6 months; signs of clinically important active inflammation of the study knee joint including redness, warmth, and/or a large bulging effusion with the loss of normal

	contour; secondary osteoarthritis of the study joint (excluded to reduce variability in the study population, which was defined as subjects with idiopathic osteoarthritis); history of acute inflammatory arthritis or pseudogout of the study joint; medical history, physical examination, or radiograph that suggested other types of arthritis, collagen vascular disease or fibromyalgia; know alcohol abuse, intravenous drug use, drug dependency, or history of significant psychiatric illness in the past 12 months; history of clinically important gastrointestinal or hepatic disease within the past 6 months; received analgesic therapy for chronic or recurrent pain conditions for indications other than osteoarthritis; received systemic corticosteroids within past 2 months, received intra-articular or periarticular corticosteroid or hyaluronan injections of the study joint within the past 6 months; received any analgesic or anti-inflammatory drugs within 5 drug half-lives plus an additional 48 hours before baseline; current use of aspirin greater than 325 mg/day, anticoagulants, psychotherapeutic agents (i.e. anticonvulsants, tranquilizers and antidepressants) with the exception of proclorperazine, statin class hypolipidaemic agents with no stable dosing regimen for at least the past 3 months, glucosamine, chondroitin sulfate, and/or shark cartilage with no stable dosing regimen for at least the past 6 months; undergoing nonpharmacological therapy requiring direct supervision or administration by a healthcare provider, such as physical therapy, massage therapy or other adjunctive therapies within the past month
Recruitment/selection of patients	People were recruited from 58 private primary care sites in the United States
Age, gender and ethnicity	Age - Mean (SD): 61.7 (10.1) years. Gender (M:F): 139:403. Ethnicity: 444 people were Caucasian, 74 were African-American, 24 were Other
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (98) or knee (443)).
Extra comments	Severity: Kellgren Lawrence grades II-III. ACR class I-III. Duration of symptoms: Not stated explicitly. At least 6 months..
Indirectness of population	No indirectness
Interventions	(n=267) Intervention 1: Paracetemol (oral) - Paracetemol. 1300mg paracetamol extended release every 8 hours for 12 weeks. Duration 12 weeks. Concurrent medication/care: People were instructed that the study medication could have contained 1300mg paracetamol per dose and that they should not exceed a maximum daily dose of 4000mg of paracetamol. A prescription of propoxyphene HCl 65mg was given as rescue medications. People were encouraged to use self-administered nonpharmacologic therapies for breakthrough pain (relaxation, stress management,

	exercise). People were not permitted to use any nonpharmacologic therapies that required the direct supervision of a healthcare provider (e.g. physical therapy) within 48 hours before returning for follow up visits. People were also not permitted to use additional analgesic medication, including rescue analgesia, within 5 drug half-lives, plus an additional 48 hours, before returning for follow-up visits. Indirectness: No indirectness (n=275) Intervention 2: Placebo. Two placebo caplets every 8 hours for 12 weeks. Duration 12 weeks. Concurrent medication/care: People were instructed that the study medication could have contained 1300mg paracetamol per dose and that they should not exceed a maximum daily dose of 4000mg of paracetamol. A prescription of propoxyphene HCl 65mg was given as rescue medications. People were encouraged to use self-administered nonpharmacologic therapies for breakthrough pain (relaxation, stress management, exercise). People were not permitted to use any nonpharmacologic therapies that required the direct supervision of a healthcare provider (e.g. physical therapy) within 48 hours before returning for follow up visits. People were also not permitted to use additional analgesic medication, including rescue analgesia, within 5 drug half-lives, plus an additional 48 hours, before returning for follow-up visits. Indirectness: No indirectness
Funding	Study funded by industry (This study was conducted, analysed and supported by McNeil Consumer Healthcare, Fort Washington, PA, USA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: Nottingham Health Profile energy subscale score at 12 weeks; Group 1: mean 2 (SD 1.27); n=267, Group 2: mean 1.72 (SD 1.28); n=275; Nottingham Health Profile energy subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Reported paracetamol: 2.00 (0.078). Reported placebo: 1.72 (0.077). Baseline paracetamol: 56.6 (38.77). Baseline placebo: 49.1 (40.06).
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -29.96 (SD 25.69); n=267, Group 2: mean -25.75 (SD 25.69); n=275;

WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and standard error. Reported paracetamol: -29.96 (1.572). Reported placebo: -25.75 (1.549). Baseline paracetamol: 78.9 (11.10). Baseline placebo: 80.8 (10.51).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others.

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -26.64 (SD 24.59); n=267, Group 2: mean -21.29 (SD 24.63); n=275; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and standard error. Reported paracetamol: -26.64 (1.505). Reported placebo: -21.29 (1.485). Baseline paracetamol: 76.1 (15.00). Baseline placebo: 75.6 (15.09).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others.

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Atrial fibrillation and congestive heart failure, chest pain, dyspnoea and angina pectoris at 12 weeks; Group 1: 4/267, Group 2: 1/275; Comments: Paracetamol: Chest pain and hypertension; atrial fibrillation; congestive cardiac failure, chest pain and dyspnoea; angina pectoris. Placebo: Atrial fibrillation and congestive heart failure.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others.

Protocol outcome 5: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Other: Increase in liver function test results >2 times the upper limit of normal at 12 weeks; Group 1: 18/267, Group 2: 2/275; Comments: Paracetamol: 1x dehydration and urosepsis, 1x renal cyst and renal mass. 7 had liver function test results exceeding 3x the upper limit of normal. 9 people had exceeded 2x but not passed 3x. Placebo: 1 exceeded 3x the upper limit of normal. 1 had exceeded 2x but not passed 3x.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention

to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others. Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 12 weeks; Group 1: 43/267, Group 2: 57/275 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, weight, OA site, patient's assessment of ostearthritis pain, Kellgren and Lawrence criteria and ACR functional class; Group 1 Number missing: 90, Reason: 267 included in the intention to treat analysis. However, 177 in the final analysis. 90 withdrew, 25 had adverse events, 42 had lack of efficacy, 6 had protocol violations, 8 withdrew consent, 4 were lost to follow up, 4 others.; Group 2 Number missing: 103, Reason: 275 included in the intention to treat analysis. However, 172 in the final analysis. 103 withdrew, 23 had adverse events, 56 had lack of efficacy, 7 protocol violations, 7 withdrew consent, 7 lost to follow up, 3 others.	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

Study	Pujalte 1980[151]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=24)
Countries and setting	Conducted in Philippines; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Established osteoarthritis of the knee (including X-ray analysis)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Outpatients with established osteoarthritis of the knee
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional inclusion criteria. No statement about glucosamine purity.
Age, gender and ethnicity	Age - Mean (range): 61.7 (45-73). Gender (M:F): 3:17. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=12) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine sulphate (2 capsules each containing 250mg, 3 times daily). Duration 6-8 weeks. Concurrent medication/care: No other analgesic, antirheumatic or anti-inflammatory drug was allowed during the observation period. Indirectness: No indirectness (n=12) Intervention 2: Placebo. Matching placebo. Duration 6-8 weeks. Concurrent medication/care: No other analgesic, antirheumatic or anti-inflammatory drug was allowed during the observation period. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO	

Protocol outcome 1: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Dizziness at 8 weeks; Group 1: 0/10, Group 2: 1/10 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex and age; Group 1 Number missing: 2, Reason: 2 people did not attend for follow up; Group 2 Number missing: 2, Reason: 2 people did not attend for follow up	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Puopolo 2007[152]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=548)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical and radiographic diagnosis of osteoarthritis of the knee or hip for at least the previous 6 months or were newly diagnosed people with clinical symptoms consistent with osteoarthritis of the study joint for at least the previous 6 months
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a clinical and radiographic diagnosis of osteoarthritis of the knee or hip for at least the previous 6 months or were newly diagnosed people with clinical symptoms consistent with osteoarthritis of the study joint for at least the previous 6 months. All eligible people met the American Rheumatism Association functional class I-III criteria and were required to have been using NSAIDs or paracetamol to treat their osteoarthritis. The primary source of pain was in the lower extremity. In cases where both knees and/or hips were affected, the most painful joint was selected for study evaluation. People who were regular users of NSAIDs (at least 25 of the last 30 days preceding enrollment) were required to have a prestudy score of less than 80mm (based on the 0-100mm visual analogue scale for the assessment of pain while walking on a flat surface); following cessation of NSAID therapy people were instructed to return to the clinic upon experiencing a flare of osteoarthritis pain. Prespecified washout periods for the various prior NSAID that were used ranged from 3 to 20 days. A sufficient flare within the washout period was defined as a patient-reported pain score of at least 40mm while the patient walked on a flat surface, and was at least 15mm greater than that recorded at the prestudy visit as well as a worsening of at least one point (0 to 5 point Likert scale) for Investigator Global Assessment of Disease Status.
Exclusion criteria	If they had medical conditions, such as recent joint injuries or rheumatologic, autoimmune, or musculoskeletal diseases that could confound or interfere with efficacy evaluations. People were excluded if they had: used intra-articular corticosteroids or hyaluronic acid injections to the study knee within the previous 3 months; used immunosuppressants within the previous 3 months; corticosteroid use

	by any systemic route; hyaluronic acid injections or intra-articular corticosteroids for any other joint in the previous month; required to take any other therapy to inhibit platelet aggregation except aspirin no more than 100mg daily.
Recruitment/selection of patients	Requires people taking NSAIDs before the study to experience a flare in symptoms after medication washout before admittance to the trial
Age, gender and ethnicity	Age - Mean (SD): 62.6 (9.5). Gender (M:F): 133:415. Ethnicity: Asian = 10, Black = 18, Hispanic American = 124, Multiracial = 131, White = 265
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American Rheumatism Association class I-III (median class II) Duration of symptoms (mean [SD]): 6.6 (7.5) years
Indirectness of population	No indirectness
Interventions	(n=437) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 800mg three times a day with matching placebo. Duration 12 weeks. Concurrent medication/care: People taking stable doses of glucosamine or chondroitin sulfate for at least 6 months prior to the study were allowed to enroll. Low-dose aspirin (no more than 100mg daily) for cardioprophylaxis was allowed. Gastroprotective agents, such as proton pump inhibitors, histamine-2 receptor antagonists, sucralfate, and misoprostol, were allowed as necessary. Paracetamol was provided as rescue medication for pain, if needed.. Indirectness: No indirectness Comments: The etoricoxib and ibuprofen groups were combined for class effect as agreed in the protocol (n=111) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: People taking stable doses of glucosamine or chondroitin sulfate for at least 6 months prior to the study were allowed to enroll. Low-dose aspirin (no more than 100mg daily) for cardioprophylaxis was allowed. Gastroprotective agents, such as proton pump inhibitors, histamine-2 receptor antagonists, sucralfate, and misoprostol, were allowed as necessary. Paracetamol was provided as rescue medication for pain, if needed.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Merck Research Laboratories)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN/ETORICOXIB versus PLACEBO	

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -26.16 (SD 23.34); n=431, Group 2: mean -16.47 (SD 21.7); n=109; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square mean change and 95% confidence intervals. Calculated SD from this. Reported etoricoxib (n=220): -28.14 (-31.24, -25.04). Reported ibuprofen (n=211): -24.10 (-27.20, -20.99). Reported placebo (n=109): -16.47 (-20.55, -12.40). Calculated SD etoricoxib: 23.5. Calculated SD ibuprofen: 23.0. Calculated SD placebo: 21.7. Baseline etoricoxib: 66.46. Baseline ibuprofen: 64.74. Baseline placebo: 64.66.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, height, weight, low dose aspirin use and baseline values of outcomes; Group 1 Number missing: 87, Reason: Ibuprofen: 15 discontinued due to lack of efficacy. 19 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 10 discontinued due to other reasons. Etoricoxib: 16 discontinued due to lack of efficacy. 11 discontinued due to clinical adverse events. 0 discontinued due to laboratory adverse events. 14 discontinued due to other reasons.; Group 2 Number missing: 32, Reason: Placebo: 21 discontinued due to lack of efficacy. 5 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 4 discontinued due to other reasons.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -21.81 (SD 22.97); n=428, Group 2: mean -13.56 (SD 21.4); n=109; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square mean change and 95% confidence intervals. Calculated SD from this. Reported etoricoxib (n=219): -23.46 (-26.78, -20.65). Reported ibuprofen (n=209): -20.09 (-23.87, -17.72). Reported placebo (n=109): -13.56 (-17.59, -9.54). Calculated SD etoricoxib: 23.1. Calculated SD ibuprofen: 22.7. Calculated SD placebo: 21.4. Baseline etoricoxib: 64.27. Baseline ibuprofen: 62.52. Baseline placebo: 64.23.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, height, weight, low dose aspirin use and baseline values of outcomes; Group 1 Number missing: 87, Reason: Ibuprofen: 15 discontinued due to lack of efficacy. 19 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 10 discontinued due to other reasons. Etoricoxib: 16 discontinued due to lack of efficacy. 11 discontinued due to clinical adverse events. 0 discontinued due to laboratory adverse events. 14 discontinued due to other reasons.; Group 2 Number missing: 32, Reason: Placebo: 21 discontinued due to lack of efficacy. 5 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 4 discontinued due to other reasons.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Discontinued due to digestive or abdominal pain adverse event at 12 weeks; Group 1: 14/437, Group 2: 4/111; Comments: Including ... Etoricoxib: Epigastric discomfort = 6, nausea = 4, dyspepsia = 7. Ibuprofen: Epigastric discomfort = 20, nausea = 9, dyspepsia = 7. Placebo: Epigastric discomfort = 2, nausea = 3, dyspepsia = 4.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports gender, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, height, weight, low dose aspirin use and baseline values of outcomes; Group 1 Number missing: 87, Reason: Ibuprofen: 15 discontinued due to lack of efficacy. 19 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 10 discontinued due to other reasons. Etoricoxib: 16 discontinued due to lack of efficacy. 11 discontinued due to clinical adverse events. 0 discontinued due to laboratory adverse events. 14 discontinued due to other reasons.; Group 2 Number missing: 32, Reason: Placebo: 21

discontinued due to lack of efficacy. 5 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 4 discontinued due to other reasons. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Discontinued due to oedema-related or hypertension-related adverse events at 12 weeks; Group 1: 6/437, Group 2: 0/111; Comments: Including ... Etoricoxib: Oedema related adverse events = 8, hypertension related adverse events = 14, congestive heart failure, pulmonary oedema or cardiac failure = 1. Ibuprofen: Oedema related adverse events = 7, hypertension related adverse events = 19, congestive heart failure, pulmonary oedema or cardiac failure = 1. Placebo: Oedema related adverse events = 2, hypertension related adverse events = 1, congestive heart failure, pulmonary oedema or cardiac failure = 0. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports gender, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, height, weight, low dose aspirin use and baseline values of outcomes; Group 1 Number missing: 87, Reason: Ibuprofen: 15 discontinued due to lack of efficacy. 19 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 10 discontinued due to other reasons. Etoricoxib: 16 discontinued due to lack of efficacy. 11 discontinued due to clinical adverse events. 0 discontinued due to laboratory adverse events. 14 discontinued due to other reasons.; Group 2 Number missing: 32, Reason: Placebo: 21 discontinued due to lack of efficacy. 5 discontinued due to clinical adverse events. 2 discontinued due to laboratory adverse events. 4 discontinued due to other reasons.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Qiu 1998[153]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=178)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 56 (10). Gender (M:F): 38:140. Ethnicity: Chinese adults
Further population details	1. Age: Mixed (Based on range 28-78). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=88) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine - two capsules 250mg glucosamine sulfate with one placebo ibuprofen tablet three times a day. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=90) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 400mg one tablet with two placebo glucosamine capsules three times a day. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus IBUPROFEN	

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Nausea, mild abdominal pain, vomiting, abdominal discomfort at 4 weeks; Group 1: 4/88, Group 2: 5/90; Comments: Glucosamine: 3 mild stomach discomfort, 1 mild nausea. Ibuprofen: 1 mild abdominal pain, 1 severe stomach discomfort, 1 moderate vomiting, 1 severe vomit, 1 severe abdominal discomfort Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age and gender; Group 1 Number missing: 1, Reason: 1 person dropped out for reasons not related to the drug; Group 2 Number missing: 9, Reason: 9 people dropped out due to drug-related adverse events Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Hypertension, oedema at 4 weeks; Group 1: 0/88, Group 2: 4/90; Comments: Ibuprofen: 1 severe hypertension, 1 mild swelling legs, 1 severe oedema of ace and legs, 1 severe oedema of eyelids and lips Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age and gender; Group 1 Number missing: 1, Reason: 1 person dropped out for reasons not related to the drug; Group 2 Number missing: 9, Reason: 9 people dropped out due to drug-related adverse events Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Haematuria at 4 weeks; Group 1: 0/88, Group 2: 1/90 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age and gender; Group 1 Number missing: 1, Reason: 1 person dropped out for reasons not related to the drug; Group 2 Number missing: 9, Reason: 9 people dropped out due to drug-related adverse events Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Sleepiness at 4 weeks; Group 1: 1/88, Group 2: 3/90; Comments: Glucosamine: 1 mild sleepiness. Ibuprofen: 1 sleepiness, 2 mild sleepiness Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age and gender; Group 1 Number missing: 1, Reason: 1 person dropped out for reasons not related to the drug; Group 2 Number missing: 9, Reason: 9 people dropped out due to drug-related adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Reed 2018[154]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=708)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a clinical diagnosis of knee or hip osteoarthritis (including grade ≥2 Kellgren-Lawrence radiographic evidence) for ≥6 months with worsening pain in the target joint requiring regular use of analgesics (≥3 days a week) during the last 3 months. Those with knee OA had to meet the modified classification criteria from the American College of Rheumatology: knee pain, radiographic osteophytes and at least one of the following: age ≥40 years, morning stiffness of <30 minutes and crepitus on motion. People with hip osteoarthirits had to have hip pain plus radiographic femoral and/or acetabular osteophytes, and radiographic joint space narrowing based on American College of Rheumatology criteria for idiopathic hip osteoarthritis.
Stratum	Other:
Subgroup analysis within study	Not applicable
Inclusion criteria	People with knee or hip osteoarthritis; age between 40-80; BMI of 18-40kg/m²; in good general health with no significant abnormalities in medical history, physical examination or laboratory evaluations. All had previously experienced pain reduction in the target joint following use of paracetamol. Women of childbearing potential were required to have negative pregnancy tests and be using reliable contraception.
Exclusion criteria	A history of surgery or major trauma to the affected joint within 6 months of screening; inflammation, secondary osteoarthritis, or history of acute inflammatory arthritis or pseudo-gout in the target joint; or inflammation in the target joint at the end of the run-in period; more than mild OA-related pain in the nontarget joints; people with a chronic pain condition or history/evidence suggestive of other types of arthritis, collagen vascular disease or fibromyalgia; a history of hepatic or renal disease; gastrointestinal surgery; current active liver or biliary disease (except Gilbert's syndrome, asymptomatic gallstones or stable chronic liver disease), including alanine aminotransferase >2 times the upper limit of normal and bilirubin >1.5 times the upper limit of normal; however, isolated bilirubin >1.5 times the upper limit of normal was acceptable if bilirubin was fractionated and direct bilirubin was <35%; current use of the following medications: analgesic therapy (including topical) for any indication other

	than osteoarthritis; anticoagulants; psychotherapeutic agents; aspirin >325mg/day; statins in doses that were not stabilized within 3 months of screening; glucosamine, chondroitin sulfate, methylsulfonylmethane or shark cartilage in doses that were not stable within 6 months of screening; oral corticosteroids within 2 months of screening; or intra-articular or periarticular corticosteroid or hyaluronan injections into the target joint within 6 months of screening.
Recruitment/selection of patients	Multisite trial
Age, gender and ethnicity	Age - Mean (SD): 60.8 (8.4). Gender (M:F): 251:425. Ethnicity: 506 were White, 143 were Black, 21 were Asian, 6 were Other/multiple
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip (57) or knee (619)).
Extra comments	Severity: Kellgren Lawrence grade 2-4. Duration of symptoms: Not stated explicitly. More then 6 months..
Indirectness of population	No indirectness
Interventions	(n=471) Intervention 1: Paracetemol (oral) - Paracetemol. Sustained release paracetamol 2x1000mg twice a day (and two placebo matched tablets three times a day) and extended release paracetamol 2x665mg three times a day (and two placebo matched tablets twice a day). Duration 12 weeks. Concurrent medication/care: People were permitted to take oral ibuprofen 2x200mg as rescue medication. With exception of the study treatment and low dose aspirin used for anti-platelet effect, all other analgesics were prohibited. Subjects were asked to wait longer than 1 hour after their last study medication dose, if possible, before taking rescue medication and to limit rescue medication use of three doses per day, separated by greater than 6 hours.. Indirectness: No indirectness Comments: Two were originally two separate groups that have been pooled together as they belong to the same class (n=237) Intervention 2: Placebo. Placebo tablets (matched to the intervention arms). Duration 12 weeks. Concurrent medication/care: People were permitted to take oral ibuprofen 2x200mg as rescue medication. With exception of the study treatment and low dose aspirin used for anti-platelet effect, all other analgesics were prohibited. Subjects were asked to wait longer than 1 hour after their last study medication dose, if possible, before taking rescue medication and to limit rescue medication use of three doses per day, separated by greater than 6 hours.. Indirectness: No indirectness

Funding	Study funded by industry (This study was funded by GlaxoSmithKline Consumer Healthcare, Parsippany, NJ, USA. Two authors (SM and KR) are employees of GlaxoSmithKline. The other author AC was an employee of GlaxoSmithKline at the time of the study.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -27.1 (SD 25.6); n=449, Group 2: mean -25.74 (SD 25.8); n=227; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and SE. Reported paracetamol SR: -28.25 (1.70). Reported paracetamol ER: -25.89 (1.71). Reported placebo: -25.74 (1.71). Baseline paracetamol SR: 69.4 (17.6). Baseline paracetamol ER: 67.9 (17.9). Baseline placebo: 72.3 (16.9).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, BMI, target joint, Kellgren-Lawrence OA score, pain rating, baseline WOMAC pain subscale score. Does not report baseline values for WOMAC physical function score.; Group 1 Number missing: 78, Reason: Modified intention to treat excluded one trial center where they had questions about their clinical practice. Includes 224 people taking SR paracetamol and 225 people taking ER paracetamol. Out of the 471 originally randomised, 30 withdrew consent, 22 had adverse events, 4 did not meet study criteria, 5 were lost to follow up, 3 had protocol violations, 14 others; Group 2 Number missing: 41, Reason: 227 were including in the modified intention to treat population. Of the 237 originally randomised, 17 withdrew consent, 8 had adverse events, 3 had protocol violations, 2 were lost to follow up, 11 others

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -25.28 (SD 25.83); n=449, Group 2: mean -23.36 (SD 25.91); n=227; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and SE. Reported paracetamol SR: -26.43 (1.71). Reported paracetamol ER: -24.13 (1.73). Reported placebo: -23.36 (1.72). Baseline values not reported.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, BMI, target joint, Kellgren-Lawrence OA score, pain rating, baseline WOMAC pain subscale score. Does not report baseline values for WOMAC physical function score.; Group 1 Number missing: 78, Reason: Modified intention to treat excluded one trial center where they had questions about their clinical practice. Includes 224 people taking SR paracetamol and 225 people taking ER paracetamol. Out of the 471 originally randomised, 30 withdrew consent, 22 had adverse events, 4 did not meet study criteria, 5 were lost to follow up, 3 had protocol violations, 14 others; Group 2 Number missing: 41, Reason: 227 were including in the modified intention to treat population. Of the 237 originally randomised, 17 withdrew consent, 8 had adverse events, 3 had protocol violations, 2 were lost to follow up, 11 others

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Hypertension at 12 weeks; Group 1: 8/470, Group 2: 1/237; Comments: Paracetamol sustained released: Hypertension = 6. Paracetamol extended release: Hypertension = 2. Placebo: Hypertension = 1.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, BMI, target joint, Kellgren-Lawrence OA score, pain rating, baseline WOMAC pain subscale score. Does not report baseline values for WOMAC physical function score.; Group 1 Number missing: 78, Reason: Modified intention to treat excluded one trial center where they had questions about their clinical practice. Includes 224 people taking SR paracetamol and 225 people taking ER paracetamol. Out of the 471 originally randomised, 30 withdrew consent, 22 had adverse events, 4 did not meet study criteria, 5 were lost to follow up, 3 had protocol violations, 14 others; Group 2 Number missing: 41, Reason: 227 were including in the modified intention to treat population. Of the 237 originally randomised, 17 withdrew consent, 8 had adverse events, 3 had protocol violations, 2 were lost to follow up, 11 others Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Other: ALT increase at 12 weeks; Group 1: 10/470, Group 2: 0/237; Comments: Paracetamol sustained released: ALT increase = 7. Paracetamol extended release: ALT increase = 4. Placebo: ALT increase = 0. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, BMI, target joint, Kellgren-Lawrence OA score, pain rating, baseline WOMAC pain subscale score. Does not report baseline values for WOMAC physical function score.; Group 1 Number missing: 78, Reason: Modified intention to treat excluded one trial center where they had questions about their clinical practice. Includes 224 people taking SR paracetamol and 225 people taking ER paracetamol. Out of the 471 originally randomised, 30 withdrew consent, 22 had adverse events, 4 did not meet study criteria, 5 were lost to follow up, 3 had protocol violations, 14 others; Group 2 Number missing: 41, Reason: 227 were including in the modified intention to treat population. Of the 237 originally randomised, 17 withdrew consent, 8 had adverse events, 3 had protocol violations, 2 were lost to follow up, 11 others Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 12 weeks; Group 1: 10/470, Group 2: 5/237; Comments: Paracetamol sustained release: 5. Paracetamol extended release: 5. Placebo: 5. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, BMI, target joint, Kellgren-Lawrence OA score, pain rating, baseline WOMAC pain subscale score. Does not report baseline values for WOMAC physical function score.; Group 1 Number missing: 78, Reason: Modified intention to treat excluded one trial center where they had questions about their clinical practice. Includes 224 people taking SR paracetamol and 225 people taking ER paracetamol. Out of the 471 originally randomised, 30 withdrew consent, 22 had adverse events, 4 did not meet study criteria, 5 were lost to follow up, 3 had protocol violations, 14 others; Group 2 Number missing: 41, Reason: 227 were including in the modified intention to treat population. Of the 237 originally randomised, 17 withdrew consent, 8 had adverse events, 3 had protocol violations, 2 were lost to follow up, 11 others	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

Study	Rindone 2000[155]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=108)
Countries and setting	Conducted in USA; Setting: Outpatient follow up (recruited from primary care)
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (2 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A history of osteoarthritis of the knee and radiographic findings consistent with the disease
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a history of osteoarthritis of the knee and radiographic findings consistent with the disease.
Exclusion criteria	People with grade 0 radiographic disease; people treated earlier with glucosamine, chondroitin or both; not ambulatory
Recruitment/selection of patients	Recruited from primary care
Age, gender and ethnicity	Age - Mean (SD): 64.5 (11.5). Gender (M:F): 93:5. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Radiographic stage 1-4 Duration of symptoms (mean [SD]): 13 (12) years.
Indirectness of population	No indirectness
Interventions	(n=54) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 500mg three times daily. Duration 8 weeks. Concurrent medication/care: People who were taking other analgesics were instructed to continue them for the duration of the study. Indirectness: No indirectness (n=54) Intervention 2: Placebo. Placebo three times daily. Duration 8 weeks. Concurrent medication/care: People who were taking other analgesics were instructed to continue them for the duration of the study. Indirectness: No indirectness
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain intensity on walking (visual analogue scale) at 8 weeks; Group 1: mean -1.4 (SD 3); n=49, Group 2: mean -1.5 (SD 2.5); n=49; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline glucosamine: 6.4 (2.5). Baseline placebo: 6.4 (2.5).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, duration of osteoarthritis, weight, pain intensity at baseline, people taking analgesics and radiographic stage; Group 1 Number missing: 5, Reason: 5 people in each group were lost to follow up; Group 2 Number missing: 5, Reason: 5 people in each group were lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Roth 2004[157]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=326)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis in at least 1 knee defined by radiological findings of deterioration and abrasion of the articular cartilage (joint space narrowing) and/or formation of new bone (osteophytes) at the joint surface of the knee (medial, lateral or patellofemoral) at an examination performed within the previous 3 months and a flare of pain after washout of stable therapy (at least 3 days per week for 1 month) consisting of an oral NSAID or paracetamol.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and nonpregnant women aged 40 to 85 years with primary osteoarthritis in at least 1 knee defined by radiological findings of deterioration and abrasion of the articular cartilage (joint space narrowing) and/or formation of new bone (osteophytes) at the joint surface of the knee (medial, lateral or patellofemoral) at an examination performed within the previous 3 months and a flare of pain after washout of stable therapy (at least 3 days per week for 1 month) consisting of an oral NSAID or paracetamol.
Exclusion criteria	Secondary arthritis related to systemic inflammatory arthritis, including rheumatoid arthritis, psoriatic arthritis, postinfectious arthritis, metabolic arthritis, and traumatic arthritis or surgical joint replacements; sensitivity to diclofenac, aspirin or any other NSAID, dimethyl sulfoxide, propylene glycol, glycerin, or ethanol; clinically active renal, hepatic, or peptic ulcer disease; a history of alcohol or other drug abuse; lactation; concomitant skin disease at the application site; corticosteroid use, including oral corticosteroid within 14 days, intramuscular corticosteroid within 30 days, intra-articular corticosteroid into the study knee within 90 days, intra-articular corticosteroid into any other joint within 30 days of study entry, or ongoing use of a topical corticosteroid at the site of application; use of a topical product, treatment or device at the application site for the relief of osteoarthritis; ongoing use of prohibited medication, including NSAID, oral analgesic, muscle relaxant or low-dose antidepressant; ongoing use of glucosamine or chondroitin sulfate sodium (unless used continuously for 90 days before study entry); intra-articular viscosupplementation (eg, hyaluronate sodium

	derivative) into the study knee in the preceding 90 days; current application for disability benefits on the basis of osteoarthritis of the knee; fibromyalgia; other painful or disabling condition affecting the knee
Recruitment/selection of patients	Requires people to have a flare in pain after washout of their previous NSAID or paracetamol therapy
Age, gender and ethnicity	Age - Mean (SD): 64.2 (10.6). Gender (M:F): 105:221. Ethnicity: White = 290, Asian = 1, Black = 30, Hispanic = 5
Further population details	1. Age: Mixed (Based on range: 41-85). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Total radiographic score (mean [SD]): 6.8 (3.7) (total maximum score = 27) Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=164) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Topical diclofenac solution, consisting of 1.5% (wt/wt) diclofenac sodium in a patented carrier containing dimethyl sulfoxide (45.5%, wt/wt), propylene glycol, glycerin, ethanol, and water. 1.3mL was applied around the affected knee (10 drops each to the front, back, medial and lateral sides), without rubbing, 4 times daily for up to 12 weeks. Duration 12 weeks. Concurrent medication/care: Rescue analgesia with paracetamol (up to four 325mg tablets per day) was permitted for residual knee or other body pain throughout the treatment period, except during the washout period before baseline and the 3 calendar days before the scheduled final assessment at week 12. Aspirin (no more than 325mg/day) was permitted for cardiovascular prophylaxis.. Indirectness: No indirectness (n=162) Intervention 2: Placebo. Topical vehicle solution, consisting of a patented carrier containing dimethyl sulfoxide (45.5%, wt/wt), propylene glycol, glycerin, ethanol, and water. 1.3mL was applied around the affected knee (10 drops each to the front, back, medial and lateral sides), without rubbing, 4 times daily for up to 12 weeks. Duration 12 weeks. Concurrent medication/care: Rescue analgesia with paracetamol (up to four 325mg tablets per day) was permitted for residual knee or other body pain throughout the treatment period, except during the washout period before baseline and the 3 calendar days before the scheduled final assessment at week 12. Aspirin (no more than 325mg/day) was permitted for cardiovascular prophylaxis.. Indirectness: No indirectness

Funding	Study funded by industry (This study was supported by Dimethaid Health Care Ltd, Markham, Ontario)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -5.9 (SD 4.7); n=163, Group 2: mean -4.3 (SD 4.4); n=159; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline diclofenac: 13.0 (3.3). Baseline placebo: 12.9 (3.4).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, radiographic score and baseline values of outcomes; Group 1 Number missing: 45, Reason: 28 lack of effect, 8 adverse events, 3 lost to follow up (1 withdrew consent, 1 could not be located, 1 was located via telephone but failed to appear for a final visit), 6 other.; Group 2 Number missing: 53, Reason: 42 lack of effect, 4 adverse events, 7 other

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -15.4 (SD 15.3); n=162, Group 2: mean -10.1 (SD 13.9); n=159; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline diclofenac: 42.0 (11.8). Baseline placebo: 41.3 (11.5).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, radiographic score and baseline values of outcomes; Group 1 Number missing: 45, Reason: 28 lack of effect, 8 adverse events, 3 lost to follow up (1 withdrew consent, 1 could not be located, 1 was located via telephone but failed to appear for a final visit), 6 other.; Group 2 Number missing: 53, Reason: 42 lack of effect, 4 adverse events, 7 other

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Melena at 12 weeks; Group 1: 0/164, Group 2: 2/162
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, radiographic score and baseline values of outcomes; Group 1 Number missing: 45, Reason: 28 lack of effect, 8 adverse events, 3 lost to follow up (1 withdrew consent, 1 could not be located, 1 was located via telephone but failed to appear for a final visit), 6 other.; Group 2 Number missing: 53, Reason: 42 lack of effect, 4 adverse events, 7 other

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Oedema at 12 weeks; Group 1: 4/164, Group 2: 2/162
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, radiographic score and baseline values of outcomes; Group 1 Number missing: 45, Reason: 28 lack of effect, 8 adverse events, 3 lost to follow up (1 withdrew consent, 1 could not be located, 1 was located via telephone but failed to appear for a final visit), 6 other.; Group 2 Number missing: 53,

Reason: 42 lack of effect, 4 adverse events, 7 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Rother 2007[159]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=397)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A minimum of a 6 months history of osteoarthritis of the knee who met two of the following 3 clinical criteria: 1) morning stiffness of <30 minutes duration, crepitus on motion and age no less than 40 years; 2) rating their pain in the index knee as no less than 3 on a five point Likert scale; 3) taking oral NSAIDs at least 3 days per week for the past 3 months or for >25 of the past 30 days.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a minimum of a 6 months history of osteoarthritis of the knee who met two of the following 3 clinical criteria: 1) morning stiffness of <30 minutes duration, crepitus on motion and age no less than 40 years; 2) rating their pain in the index knee as no less than 3 on a five point Likert scale; 3) taking oral NSAIDs at least 3 days per week for the past 3 months or for >25 of the past 30 days. They also had to meet three osteoarthritis flare criteria: 1) pain in the index knee on walking no less than 40mm on a visual analogue scale, 2) increased by no less than 15mm compared with pain on prestudy treatment (screening); 3) patient global assessment score for osteoarthritis of 3-5 and at least one grade increase from screening.
Exclusion criteria	Grade 1 or grade 4 severity of the index knee based on the Kellgren and Lawrence radiographic criteria; intra-articular injections or arthroscopy of the index knee within 3 months before screening; signs of any clinically important inflammation of the index knee; crystalline-induced synovitis in the index knee; a history, physical examination or radiography results suggestive of acute inflammatory arthritis, rheumatoid arthritis, psoriatic arthritis, septic arthritis, gout, pseudogout, fibromyalgia, lupus erythematosus, or other types of inflammatory arthritis of the index knee
Recruitment/selection of patients	They had to meet three osteoarthritis flare criteria: 1) pain in the index knee on walking no less than 40mm on a visual analogue scale, 2) increased by no less than 15mm compared with pain on prestudy treatment (screening); 3) patient global assessment score for osteoarthritis of 3-5 and at least one grade increase from screening.
Age, gender and ethnicity	Age - Mean (SD): 62.8 (9.8). Gender (M:F): 160:237. Ethnicity: Not stated

Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 6 months..
Indirectness of population	No indirectness
Interventions	(n=138) Intervention 1: NSAID gels (topical - local) - Other. 110mg epicutaneous ketoprofen in 4.8 grams Transfersome (a semi-solid formulation, IDEA-033) and oral placebo every 12 hours for 6 weeks. Duration 6 weeks. Concurrent medication/care: People could take up to 2000mg paracetamol per day as rescue medication for knee pain for 3 days in any week, apart from the 48 hours preceding a study visit. Indirectness: No indirectness (n=132) Intervention 2: NSAIDs - Celecoxib. Celecoxib 100mg orally and placebo gel twice a day. Duration 6 weeks. Concurrent medication/care: People could take up to 2000mg paracetamol per day as rescue medication for knee pain for 3 days in any week, apart from the 48 hours preceding a study visit. Indirectness: No indirectness (n=127) Intervention 3: Placebo. Matching oral and topical placebo twice a day. Duration 6 weeks. Concurrent medication/care: People could take up to 2000mg paracetamol per day as rescue medication for knee pain for 3 days in any week, apart from the 48 hours preceding a study visit. Indirectness: No indirectness
Funding	Study funded by industry (IDEA AG (Germany) and McNeil Consumer and Speciality Pharmaceuticals (USA) sponsored the study and carried out on-site monitoring of all participants. The sponsors had an opportunity to comment on the manuscript before submission, but the final version was the sole responsibility of the authors.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus CELECOXIB

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -19.4 (SD 21.2); n=138, Group 2: mean -20.7 (SD 22.7); n=132; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported ketoprofen: -19.4 (21.2). Reported celecoxib: -20.7 (22.7). Baseline ketoprofen: 55.1 (18.4). Baseline celecoxib: 56.1 (18.6).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -16 (SD 20.3); n=138, Group 2: mean -18.1 (SD 22.5); n=132; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported ketoprofen: -16.0 (20.3). Reported celecoxib: -18.1 (22.5). Baseline ketoprofen: 53.8 (20.4). Baseline celecoxib: 54.6 (21.0).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders, any adverse events at 6 weeks; Group 1: 13/138, Group 2: 18/132; Comments: Including ... Ketoprofen: upper abdominal pain = 2, constipation = 2, diarrhoea = 1, dyspepsia = 1, flatulence = 0, gastritis = 3, nausea = 2. Celecoxib: upper abdominal pain = 4, constipation = 0, diarrhoea = 2, dyspepsia = 4, flatulence = 2, gastritis = 0, nausea = 3.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Depression at 6 weeks; Group 1: 0/138, Group 2: 3/132
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -19.4 (SD 21.2); n=138, Group 2: mean -12.4 (SD 20.8); n=127; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported ketoprofen: -19.4 (21.2). Reported placebo: -12.4 (20.8). Baseline ketoprofen: 55.1 (18.4). Baseline placebo: 59.9 (17.3).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.

Protocol outcome 2: Physical function at ≤3- or >3- months

- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -16 (SD 20.3); n=138, Group 2: mean -12.3 (SD 19.2); n=127; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported ketoprofen: -16.0 (20.3). Reported placebo: -12.3 (19.2). Baseline ketoprofen: 53.8 (20.4). Baseline placebo: 58.9 (19.6).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders, any adverse events at 6 weeks; Group 1: 13/138, Group 2: 12/127; Comments: Including ... Ketoprofen: upper abdominal pain = 2, constipation = 2, diarrhoea = 1, dyspepsia = 1, flatulence = 0, gastritis = 3, nausea = 2. Placebo: upper abdominal pain = 3, constipation = 1, diarrhoea = 0, dyspepsia = 1, flatulence = 0, gastritis = 3, nausea = 2.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.

Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Depression at 6 weeks; Group 1: 0/138, Group 2: 0/127
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 25, Reason: Ketoprofen: 23 adverse events, 1 lack of efficacy, 1 lost to follow up.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -20.7 (SD 22.7); n=132, Group 2: mean -12.4 (SD 20.8); n=127; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported celecoxib: -20.7 (22.7). Reported placebo: -12.4 (20.8). Baseline celecoxib: 56.1 (18.6). Baseline placebo: 59.9 (17.3).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -18.1 (SD 22.5); n=132, Group 2: mean -12.3 (SD 19.2); n=127; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Extracted mean (SD) change in pain score. Reported celecoxib: -18.1

(22.5). Reported placebo: -12.3 (19.2). Baseline celecoxib: 54.6 (21.0). Baseline placebo: 58.9 (19.6). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other. Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal disorders, any adverse events at 6 weeks; Group 1: 18/132, Group 2: 12/127; Comments: Including ... Celecoxib: upper abdominal pain = 4, constipation = 0, diarrhoea = 2, dyspepsia = 4, flatulence = 2, gastritis = 0, nausea = 3. Placebo: upper abdominal pain = 3, constipation = 1, diarrhoea = 0, dyspepsia = 1, flatulence = 0, gastritis = 3, nausea = 2. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other. Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Depression at 6 weeks; Group 1: 3/132, Group 2: 0/127 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, and baseline values of outcomes; Group 1 Number missing: 23, Reason: Celecoxib: 18 adverse events, 3 lack of efficacy, 1 subject request, 1 other.; Group 2 Number missing: 25, Reason: Placebo: 20 adverse events, 3 lack of efficacy, 1 protocol violation, 1 other.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Rother 2013[158]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=555)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A primary diagnosis of functional class I-III osteoarthritis of the knee according to the American College of Rheumatology clinical classification criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged >45 years with a primary diagnosis of functional class I-III osteoarthritis of the knee. People had to be able to identify a predominantly painful knee and to experience moderate pain in the index knee when walking on a flat surface defined as a score of ≥4 on question 1 of the WOMAC OA index, and a total average WOMAC pain subscale score of <7 at B1 and B2. Females of child-bearing potential were to use suitable birth control methods and to have a negative pregnancy test at screening and B2.
Exclusion criteria	The difference between pain rating (determined by question 1 of the WOMAC 11-point NRS) or the difference in the total average pain score between B1 and B2 were >2 (range 0-10); skin lesions or dermatologic diseases in the treatment area; extreme obesity (body mass index >37 kg/m²; symptomatic ipsilateral hip OA; inflammatory arthritis; malignancy within the previous 2 years; epilepsy; schizophrenia; any pain condition requiring the chronic use of pain medication; hypersensitivity or allergy to NSAID, including ketoprofen, or prexisting asthma or bronchospasm following NSAID use; people with known GI and CV risk factors for NSAID; intraarticular injections of hyaluronic acid products in the index knee; arthroscopy of the index knee; use of tricyclic anti-depressants within 3 months prior to or during the study; use of oral, inhaled or parenteral corticosteroids within 2 months prior to or during the study; intraarticular injections of corticosteroids in the index knee within 1 month prior to or during the study; receipt of any investigational product within 30 days of the screening visit; participation in any previous clinical trial of ketoprofen in transfersome gel.

Recruitment/selection of patients	Everyone was assessed at baseline visit 1, then had a washout period of more than 5 days to remove any previous medication before a second baseline visit. Excludes people who had a difference of >2 in their baseline visit 1 and 2 scores.
Age, gender and ethnicity	Age - Mean (SD): 62.2 (10.4). Gender (M:F): 209:346. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional Class I-III Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=274) Intervention 1: NSAID gels (topical - local) - Other. Ketoprofen gel - 100mg in 4.4g transfersome gel (IDEA-033) - administered topically twice daily. Duration 12 weeks. Concurrent medication/care: Rescue medication (500mg paracetamol, up to 4 times per day, total 2g) was permitted for the treatment of intermittent pain, but not within 24 hours of the next study visit or between B1 and B2. People who required ≥2g/day of rescue or other analgesic medication for >3 consecutive days were considered to be treatment failures and were withdrawn from the study. Indirectness: No indirectness (n=281) Intervention 2: Placebo. 4.4g ketoprofen-free vehicle only (transfersome gel only) - administered topically twice daily. Duration 12 weeks. Concurrent medication/care: Rescue medication (500mg paracetamol, up to 4 times per day, total 2g) was permitted for the treatment of intermittent pain, but not within 24 hours of the next study visit or between B1 and B2. People who required ≥2g/day of rescue or other analgesic medication for >3 consecutive days were considered to be treatment failures and were withdrawn from the study. Indirectness: No indirectness
Funding	Study funded by industry (Supported by IDEA AGE, Germany (sponsor of the clinical trial))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: KETOPROFEN GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -1.98 (SD 1.97); n=274, Group 2: mean -2.33 (SD 2.12); n=281; WOMAC pain subscale 0-10 Top=High is poor outcome; Comments: Baseline ketoprofen: 5.2 (1.0). Baseline placebo: 5.3 (1.0).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean BMI, mean baseline scores and

analgesic use at screening; Group 1 Number missing: 54, Reason: 20 withdrew due to adverse events, 16 lack of efficacy, 10 withdrawal of consent, 15 administrative (including personal reasons, use of excluded medications and non-compliance), 3 lost to follow up; Group 2 Number missing: 51, Reason: 11 withdrawals after adverse events, 18 lack of efficacy, 8 withdrawal of consent, 15 administrative, 6 lost to follow up

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -2.02 (SD 2.07); n=274, Group 2: mean -2.32 (SD 2.23); n=281; WOMAC physical function subscale 0-10 Top=High is poor outcome; Comments: Baseline ketoprofen: 5.4 (1.2). Baseline placebo: 5.4 (1.2).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean BMI, mean baseline scores and analgesic use at screening; Group 1 Number missing: 54, Reason: 20 withdrew due to adverse events, 16 lack of efficacy, 10 withdrawal of consent, 15 administrative (including personal reasons, use of excluded medications and non-compliance), 3 lost to follow up; Group 2 Number missing: 51, Reason: 11 withdrawals after adverse events, 18 lack of efficacy, 8 withdrawal of consent, 15 administrative, 6 lost to follow up

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Flatulence, nausea, haemorrhoids and upset stomach at 12 weeks; Group 1: 2/274, Group 2: 2/281; Comments: Ketoprofen: Haemorrhoids = 1, upset stomach = 1. Placebo: Flatulence = 1, nausea = 1
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean BMI, mean baseline scores and analgesic use at screening; Group 1 Number missing: 54, Reason: 20 withdrew due to adverse events, 16 lack of efficacy, 10 withdrawal of consent, 15 administrative (including personal reasons, use of excluded medications and non-compliance), 3 lost to follow up; Group 2 Number missing: 51, Reason: 11 withdrawals after adverse events, 18 lack of efficacy, 8 withdrawal of consent, 15 administrative, 6 lost to follow up

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cardiac or vascular adverse events at 12 weeks; Group 1: 0/274, Group 2: 0/281; Comments: Reports that there were no cardiac or vascular adverse events
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, mean BMI, mean baseline scores and analgesic use at screening; Group 1 Number missing: 54, Reason: 20 withdrew due to adverse events, 16 lack of efficacy, 10 withdrawal of consent, 15 administrative (including personal reasons, use of excluded medications and non-compliance), 3 lost to follow up; Group 2 Number missing: 51, Reason: 11 withdrawals after adverse events, 18 lack of efficacy, 8 withdrawal of consent, 15 administrative, 6 lost to follow up

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Rovensky 2001[160]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Germany, Slovakia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 week
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis, unilateral or bilateral that is chronic and decompensated (painful, but nonactivated and without effusion or swelling) diagnosed according to the American College of Rheumatology classification criteria: knee pain for most days (>15 days) of the preceding month and radiographic osteophytes, grade 2 or 3 on the Kellgren and Lawrence osteoarthritis severity score, a pain on motion score of 40mm on a 100-mm visual analogue scale, total score of 5 and 13 on the Lequesne algofunctional index
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and females aged 40 to 75 years with primary knee osteoarthritis (as defined above) who provided written informed consent.
Exclusion criteria	Secondary osteoarthritis; obesity (body mass index ≥30 kg/m²; chronic painful disease of the hip or ankle joint; allergic diathesis, bronchial asthma or known hypersensitivity to nonsteroidal antiinflammatory drugs (NSAIDs); eczematous skin eruption; people undergoing physiotherapy
Recruitment/selection of patients	Some people were drug naive. If washout was required this occurred over 1-60 days with no specific flare criteria.
Age, gender and ethnicity	Age - Mean (SD): 63.4 (8.3). Gender (M:F): 26:74. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated explicitly. At least 1 month.
Indirectness of population	No indirectness
Interventions	(n=50) Intervention 1: NSAID gels (topical - local) - Ibuprofen gel. 5% ibuprofen cream - apply a 10cm strip of cream on the more painful knee and apply 3 times daily (corresponding to 3x 4g of cream and 3x 200mg ibuprofen). Duration 1 week. Concurrent medication/care: Any concomitant treatment with other topical,

	intraarticular or systemic steroidal or nonsteroidal antirheumatics, analgesics or disease-modifying antirheumatic drugs was not allowed. However, people were allowed to receive any medical treatment required, even if this led to their exclusion from the study. During the washout period, peripherally acting oral analgesics, such as paracetamol, were allowed as escape medication up to 2 days before the start of the study treatment. Indirectness: No indirectness (n=50) Intervention 2: Placebo. Placebo cream - apply a 10cm strip of cream on the more painful knee and apply 3 times daily (corresponding to 3x 4g of cream). Duration 1 week. Concurrent medication/care: Any concomitant treatment with other topical, intraarticular or systemic steroidal or nonsteroidal antirheumatics, analgesics or disease-modifying antirheumatic drugs was not allowed. However, people were allowed to receive any medical treatment required, even if this led to their exclusion from the study. During the washout period, peripherally acting oral analgesics, such as paracetamol, were allowed as escape medication up to 2 days before the start of the study treatment. Indirectness: No indirectness
Funding	Other author(s) funded by industry (Lenhard G. is an employee of Dr. med. G. Lenhard and Partner GmbH, Overath, Germany. Vogtle-Junkert and Schreyger, F. are employees of Dolorgiet Pharmaceuticals, St. Augustin/Bonn, Germany)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain on motion at 1 week (8 days); Group 1: mean 41.66 (SD 15.32); n=50, Group 2: mean 52 (SD 16.78); n=50; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 65.90 (9.70). Baseline placebo: 64.38 (9.45).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI, Kellgren Lawrence severity score and baseline values of outcomes; Group 1 Number missing: 0, Reason: 0 discontinued.; Group 2 Number missing: 4, Reason: 3 discontinued for use of nonallowed drug. 1 the reasons was not given for.

Protocol outcome 2: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Cystopyelitis at 1 week (8 days); Group 1: 1/50, Group 2: 0/50
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI, Kellgren Lawrence severity score and baseline values of outcomes; Group 1 Number missing: 0, Reason: 0 discontinued.; Group 2 Number missing: 4, Reason: 3 discontinued for use of nonallowed drug. 1 the reasons was not given for.

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Rozendaal 2008[161] (Rozendaal 2009[162])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=222)
Countries and setting	Conducted in Netherlands; Setting: Primary care follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who met the American College of Rheumatology clinical criteria for hip osteoarthritis
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People who met the American College of Rheumatology criteria for hip osteoarthritis
Exclusion criteria	People who had undergone or were awaiting hip replacement surgery; Kellgren and Lawrence score of 4; renal disease; liver disease; diabetes mellitus; a disabling comorbid condition that would make visits to the research center impossible; people already receiving glucosamine; those unable to fill out Dutch questionnaires
Recruitment/selection of patients	No additional criteria for inclusion
Age, gender and ethnicity	Age - Mean (SD): 63.4 (9.0). Gender (M:F): 68:154. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Kellgren Lawrence grade 1-4 Duration of symptoms: <1->3 years - the majority were for >3 years
Indirectness of population	No indirectness
Interventions	(n=111) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 1500mg (administered once daily as two 750mg tablets). Contained 200mg of D-glucosamine sulfate 2-potassium chloride, which results in a net content of 1500mg of glucosamine sulfate per 2 pills.. Duration 2 years. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=111) Intervention 2: Placebo. Matching placebo once a day. Duration 2 years. Concurrent medication/care: No additional information. Indirectness: No indirectness

Funding	Academic or government funding (Supported by the Erasmus Medical Center-Breedtestrategic program)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hip: WOMAC pain subscale at 12 weeks (3 months); Group 1: mean -2.5 (SD 19.2); n=111, Group 2: mean -1.79 (SD 16.2); n=111; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 35.9 (23.0). Baseline placebo: 32.4 (23.2).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, osteoarthritis type, Kellgren and Lawrence score, mean minimum joint space width, WOMAC scores, mean pain in the past week, pain medication use and assumed medication type.; Group 1 Number missing: 7, Reason: 2 deceased, 1 lack of efficacy, 2 adverse events, 2 personal circumstances; Group 2 Number missing: 8, Reason: 1 deceased, 2 unknown reason, 2 lack of efficacy, 2 bilateral total hip arthroplasty, 1 adverse event
- Actual outcome for Hip: WOMAC pain subscale at 2 years; Group 1: mean -1.9 (SD 16.9); n=111, Group 2: mean -0.3 (SD 16.9); n=111; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Converted into standard deviation. Reported glucosamine: -1.90 (1.6). Reported placebo: -0.3 (1.6). Baseline glucosamine: 35.9 (23.0). Baseline placebo: 32.4 (23.2).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, osteoarthritis type, Kellgren and Lawrence score, mean minimum joint space width, WOMAC scores, mean pain in the past week, pain medication use and assumed medication type.; Group 1 Number missing: 7, Reason: 2 deceased, 1 lack of efficacy, 2 adverse events, 2 personal circumstances; Group 2 Number missing: 8, Reason: 1 deceased, 2 unknown reason, 2 lack of efficacy, 2 bilateral total hip arthroplasty, 1 adverse event

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hip: WOMAC function subscale at 12 weeks (3 months); Group 1: mean -3.29 (SD 14.9); n=111, Group 2: mean -1.08 (SD 12.7); n=111; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Baseline glucosamine: 36.0 (24.1). Baseline placebo: 34.1 (21.7).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, osteoarthritis type, Kellgren and Lawrence score, mean minimum joint space width, WOMAC scores, mean pain in the past week, pain medication use and assumed medication type.; Group 1 Number missing: 7, Reason: 2 deceased, 1 lack of efficacy, 2 adverse events, 2 personal circumstances; Group 2 Number missing: 8, Reason: 1 deceased, 2 unknown reason, 2 lack of efficacy, 2 bilateral total hip arthroplasty, 1 adverse event
- Actual outcome for Hip: WOMAC function subscale at 2 years; Group 1: mean -1.69 (SD 13.7); n=111, Group 2: mean 0.38 (SD 13.7); n=111; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Converted into standard deviation. Reported glucosamine: -1.69 (1.3). Reported placebo: +0.38 (1.3). Baseline glucosamine: 36.0 (24.1). Baseline placebo: 34.1 (21.7).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, duration of symptoms, osteoarthritis type, Kellgren and Lawrence score, mean minimum joint space width, WOMAC scores, mean pain in the past week, pain medication use and assumed medication type.; Group 1 Number missing: 7, Reason: 2 deceased, 1 lack of efficacy, 2 adverse events, 2 personal circumstances; Group 2 Number

missing: 8, Reason: 1 deceased, 2 unknown reason, 2 lack of efficacy, 2 bilateral total hip arthroplasty, 1 adverse event	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Sanda 1983[163]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=58)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed by definite degenerative joint disease of the hip confirmed by roentgenological and clinical criteria with X-rays obtained within 3 months prior to admission showing definite osteophyte formation and one or more of the following: lipping of marginal bone, narrowing of the joint space, sharpened articular margin, damaged, thickened or eburnated subchondral bone, or a bone cyst.
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed as having definite degenerative joint disease of the hip confirmed by roentgenological and clinical criteria confirmed with x-ray features. In addition, one or more of the following: limitation of movement, night pain, tenderness on pressure, morning stiffness. A positive response to any NSAID used for treatment of osteoarthritis in the past was also required.
Exclusion criteria	People who did not have a flare during the 2 week washout period
Recruitment/selection of patients	Multicenter study conducted by 12 investigators
Age, gender and ethnicity	Age - Mean (range): 59 (33-70). Gender (M:F): 13:45. Ethnicity: 56 White, 2 Black
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Not stated Duration of symptoms: Not stated. Flare design - dependent on the participant having a flare during the two weeks maximum washout period at the start of the study
Indirectness of population	No indirectness
Interventions	(n=42) Intervention 1: NSAIDs - High-dose aspirin. Aspirin 2400-4800 mg/day in four doses or Etodolac 100-400mg/day in two doses with two matching placebo tablets (to make up medication administration at four times a day). Duration 12 weeks. Concurrent medication/care: Non-narcotic, analgesic paracetamol (650mg four times daily PRN) was permitted only during the washout period. Indirectness: No

	indirectness Comments: High-dose aspirin and etodolac groups were combined for class effect as agreed in the protocol (n=16) Intervention 2: Placebo. Placebo four times a day. Duration 12 weeks. Concurrent medication/care: Non-narcotic, analgesic paracetamol (650mg four times daily PRN) was permitted only during the washout period. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: HIGH-DOSE ASPIRIN/ETODOLAC versus PLACEBO Protocol outcome 1: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hip: All central nervous system complaints, dizziness, tinnitus, hearing loss at 12 weeks; Group 1: 12/42, Group 2: 4/16; Comments: Etodolac: CNS all complaints = 2. Aspirin: CNS all complaints = 10 with 5 being due to dizziness, tinnitus = 4, hearing loss = 3. Placebo: CNS all complaints = 4. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, duration of osteoarthritis and days taken to enter a flare; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Sandelin 1997[164]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=281)
Countries and setting	Conducted in Finland, Sweden; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiologically confirmed osteoarthritis (including osteophytes) of one or both knees and with pain symptoms for most days of the prior month to such an extent that administration of analgesics was needed
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with radiologically confirmed osteoarthritis of one or both knees, and with pain symptoms for most days of the prior month to such an extent that administration of analgesics was needed
Exclusion criteria	People with severe osteoarthritis, ISK-score >14; other known bone and/or joint diseases which might have interfered with the results (e.g. chondromatosis, gout, arthritis psoriatica, infectious or postinfectious forms of arthritis, rheumatoid arthritis, chondrocalcinosis); known neurological or mechanical abnormality in the affected limb with significant functional impairment; implant arthroplasty of any kind in the lower limbs; symptoms of osteoarthritis of the hip that could influence the ISK-score; injections or systemic treatment with corticosteroids; administration of disease modifying drugs or any intraarticularly administered drugs in the previous 4 months; intake of paracetamol, acetylsalicylic acid or any NSAID with in the last 2 days (4 days for NSAIDs with long half-lives as piroxicam and tenoxicam) before enrollment examination as well as use of paracetamol, ASA or any NSAID on a regular basis (more than 6 tablets weekly) or local administration of topical NSAIDs or corticosteroids on the affected knee during the last 7 days before enrollment examination; people who begun physiotherapy or use of orthotic devices within the previous 7 days (after this time period people were recommended to continue these treatments if included in the study); known haematological blood disorders; known hypersensitivity to ASA or other NSAIDs; clinically relevant dysfunction of liver or kidneys; known gastric and/or duodenal ulcers; pregnancy or breast-feeding; known contraindication or interaction of diclofenac.
Recruitment/selection of patients	Multicentre trial

Age, gender and ethnicity	Age - Mean (SD): 61 (8.1). Gender (M:F): 92:189. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=126) Intervention 1: NSAID gels (topical - local) - Other. Eltenac gel 1%, 3g (=300mg eltenac) applied 3 times daily combined with 1 placebo tablet twice daily. Duration 4 weeks. Concurrent medication/care: No additional treatment. Indirectness: No indirectness Comments: Elternac is not licensed for use in the UK so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=82) Intervention 2: NSAIDs - Diclofenac. Diclofenac 50mg twice daily combined with placebo gel 3g applied three times daily. Duration 4 weeks. Concurrent medication/care: No additional treatment. Indirectness: No indirectness (n=82) Intervention 3: Placebo. Placebo gel, 3g applied three times daily combined with 1 placebo tablet twice daily. Duration 4 weeks. Concurrent medication/care: No additional treatment. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Patient's assessment of overall pain (VAS) at 4 weeks; Group 1: mean 30 (SD 19.2); n=82, Group 2: mean 32 (SD 24.1); n=82; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 52 (20.4). Baseline placebo: 53 (22.2).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, BMI, obesity, bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 4, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 4 were from the diclofenac group.; Group 2 Number missing: 3, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 3 were from the placebo group.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: GI reactions at 4 weeks; Group 1: 11/82, Group 2: 6/82; Comments: Including diarrhoea, nausea and abdominal pain Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, BMI, obesity, bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 4, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 4 were from the diclofenac group.; Group 2 Number missing: 3, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 3 were from the placebo group. Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: CNS reactions at 4 weeks; Group 1: 6/82, Group 2: 4/82 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, BMI, obesity, bilateral osteoarthritis and baseline values of outcomes; Group 1 Number missing: 4, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 4 were from the diclofenac group.; Group 2 Number missing: 3, Reason: Overall, of the 290 people randomised, 3 discontinued for non-medical reasons, 6 for joint disease other than osteoarthritis that were regarded unsuited for treatment with topical NSAIDs alone. 3 were from the placebo group.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Schiff 1996[167]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=347)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee diagnosed at least 6 months prior to screening with at least a physician and self-baseline ratings of their condition as less than 9 on an 11 point scale with a flare of disease activity within 2 weeks of discontinuing previous NSAID therapy and a history of clinical response to an NSAID
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People between the ages of 18 and 80 years, with osteoarthritis of the knee diagnosed at least 6 months prior to screening with at least a physician and self-baseline ratings of their condition as less than 9 on an 11 point scale with a flare of disease activity within 2 weeks of discontinuing previous NSAID therapy and a history of clinical response to an NSAID
Exclusion criteria	Women of childbearing age who were either pregnant, lactating or not using an adequate method of birth control; those who had a history of significant gastrointestinal haemorrhage within the past 2 years; had a haemoglobin value <10g/dL or positive occult blood in the stool; if within 48 hours of study entry they had used histamine-2 receptor blocking agents, antacids, sucralfate, anticholinergics, misoprostol, or other drugs known to mask or affect GI tolerability; required drugs that could confound quantification of analgesia, such as antidepressants or analgesics; those who received any intra-articular steroid within 6 months prior to study entry; those who had a known or suspected allergy or sensitivity to naproxen, aspirin or any other NSAID
Recruitment/selection of patients	People had to have had a recent osteoarthritis flare and a previous response to NSAIDs to be included
Age, gender and ethnicity	Age - Other: Mean (SE): Naprelan = 63.2 (0.84). Naprosyn = 64.1 (0.93). Placebo = 64.2 (0.87).. Gender (M:F): 109:238. Ethnicity: White = 288, Black = 40, Other = 19
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 6 months. Followed by an open phase extension study that was not included in the analysis
Indirectness of population	No indirectness
Interventions	(n=231) Intervention 1: NSAIDs - Naproxen. Naprelan (long-acting naproxen 2x500mg once in the morning and 1 placebo tablet in the morning and one in the evening) or Naprosyn (500mg in the morning and the evening with 2 placebo tablets in the morning). Duration 12 weeks. Concurrent medication/care: They were allowed to use commercially available paracetamol 325mg for rescue analgesia. Indirectness: No indirectness Comments: These two groups were combined in the analysis due to class effect as agreed in the protocol (n=116) Intervention 2: Placebo. 3 placebo tablets in the morning and 1 in the evening. Duration 12 weeks. Concurrent medication/care: They were allowed to use commercially available paracetamol 325mg for rescue analgesia. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Hypertension at 12 weeks; Group 1: 9/231, Group 2: 3/116; Comments: Napelan: 3. Naprosyn: 6. Placebo: 3.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, sex and race; Group 1 Number missing: , Reason: 102 out of 347 failed to complete 12 weeks of the double blind treatment. The most frequent reasons for termination were adverse events (12 naprelan, 15 naprosyn, 2 placebo) and treatment failure (11 naprelan, 11 naprosyn, 32 placebo).; Group 2 Number missing: , Reason: 102 out of 347 failed to complete 12 weeks of the double blind treatment. The most frequent reasons for termination were adverse events (12 naprelan, 15 naprosyn, 2 placebo) and treatment failure (11 naprelan, 11 naprosyn, 32 placebo).

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache at 12 weeks; Group 1: 50/231, Group 2: 21/116; Comments: Naprelan: 22. Naprosyn: 28. Placebo: 21.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, sex and race; Group 1 Number missing: , Reason: 102 out of 347 failed to complete 12 weeks of the double blind treatment. The most frequent reasons for termination were adverse events (12 naprelan, 15 naprosyn, 2 placebo) and treatment failure (11 naprelan, 11 naprosyn, 32 placebo).; Group 2 Number missing: , Reason: 102

out of 347 failed to complete 12 weeks of the double blind treatment. The most frequent reasons for termination were adverse events (12 naprelan, 15 naprosyn, 2 placebo) and treatment failure (11 naprelan, 11 naprosyn, 32 placebo).	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Schmitt 1999[168]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=393)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Activated osteoarthritis of the knee and/or hip with clinical features, fulfilling diagnostic criteria, functional assessment and radiological classifications of osteoarthritis according to Lequesne
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Female and male outpatients with activated osteoarthritis of the knee who gave informed consent. After a washout period of previous medication they needed to score at least 40 mm on a visual analogue scale
Exclusion criteria	People who had undergone intra-articular infiltrations or punctions, synovectomy, arthrodesis or surgery in the proceeding months; those with major functional disorders that could rapidly necessitate an operative indication; necrosis of the femoral head or condyle; arthropathy associated with Paget's disease; haemophilic arthropathy; chondrocalcinosis; ochronosis; haemochromatosis; isolated or predominant inflammatory arthritis (rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis); tuberculous arthritis; Charcot's hip/knee joint; villonodular synovitis; synovial chondromatosis; pregnant or nursing women or women capable of bearing without effective contraception; clinically relevant psychiatric disorders (ICD-10); malignant diseases or history of malignant diseases; untreated hypertension (WHO III or IV); chronic heart failure (NYHA III or IV); clinically relevant renal (serum creatinine above two times upper limit of normal); diabetes mellitus; hepatic insufficiency (SGPT more than 1.5 times above upper limit of normal); NSAID-induced porphyria; chronic pancreatitis; asthma or hay fever; active skin or mucosa disorders; allergy to NSAIDs and inhibitors of prostaglandin synthesis; alcohol or drug abuse; participation in another clinical trial; active peptic ulcer or documented history of peptic ulcer within the last 3 years; gastric bleeding; history of gastritis or active gastritis; occult changes in the blood count
Recruitment/selection of patients	People had to have pain over a certain level (40mm on a VAS) after a washout of previous medication

Age, gender and ethnicity	Age - Mean (SD): 61 (9). Gender (M:F): 64:329. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=337) Intervention 1: NSAIDs - Diclofenac. Diclofenac sodium 150mg dual release capsule (containing 100mg of diclofenac sodium sustained in release pellets and 50mg diclofenac sodium in enteric coated pellets) every morning and placebo capsule 6 and 12 hours thereafter, or diclofenac sodium 75mg dual release capsule (containing 50mg of diclofenac sodium in sustained release pellets and 25mg diclofenac sodium in enteric coated pellets) every morning and a placebo capsule 6 and 12 hours thereafter or diclofenac sodium 50mg enteric coated tablet in a capsule every morning and 6 and 12 hours thereafter. Duration 12 weeks. Concurrent medication/care: The application of corticosteroids and NSAIDs administered topically or in infiltration form as well as all anti-inflammatory and analgesic drugs were not permitted during the active treatment phase. During the washout period, the administration of oral NSAIDs was not allowed. However, paracetamol was allowed until one day before the baseline visit. Furthermore the following treatment was contraindicated: lithium or digitalis drugs, diuretics, anticoagulants, methotrexate, phenytoin, acetylsalicylic acid, ticlopidine and cyclosporine. Indirectness: No indirectness Comments: The three diclofenac groups were combined due to class effect as agreed in the protocol (n=56) Intervention 2: Placebo. A placebo capsule every morning and 6 and 12 hours thereafter. Duration 12 weeks. Concurrent medication/care: The application of corticosteroids and NSAIDs administered topically or in infiltration form as well as all anti-inflammatory and analgesic drugs were not permitted during the active treatment phase. During the washout period, the administration of oral NSAIDs was not allowed. However, paracetamol was allowed until one day before the baseline visit. Furthermore the following treatment was contraindicated: lithium or digitalis drugs, diuretics, anticoagulants, methotrexate, phenytoin, acetylsalicylic acid, ticlopidine and cyclosporine. Indirectness: No indirectness

Funding	Principal author funded by industry (All authors were either employees of Klinge Pharma GmbH, Munich/Germany or Chrysalis Clinical Pharmacology Services GmbH, Duesseldorf/Germany)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Area under the curve visual analogue scale calculated from the 24 hour visual analogue scale profile at 12 weeks; Group 1: mean 31.6 (SD 24.2); n=337, Group 2: mean 36.3 (SD 24.2); n=56; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports area under the curve values with the units mm/h. Assuming their units are incorrect and should be mmh, these values have been multiplied by 24 hours to determine the visual analogue scale score. Reported DRC150: 791.4 (612.4). Reported DRC75: 779.7 (576.8). Reported EC50: 698.2 (546.9). Reported placebo: 871.4 (579.8). Calculated DRC150: 33.0 (25.5). Calculated DRC75: 32.5 (24.0). Calculated EC50: 29.1 (22.8). Calculated placebo: 36.3 (24.2). Reported baseline DRC150: 1500.3 (322.1). Reported baseline DRC75: 1524.2 (339.3). Reported baseline EC50: 1513.6 (325.4). Reported baseline placebo: 1521.8 (308.5). Calculated baseline DRC150: 62.5 (13.4). Calculated baseline DRC75: 63.5 (14.1). Calculated baseline EC50: 63.1 (13.6). Calculated baseline placebo: 63.4 (12.9).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Downgraded for outcome reporting due to the value for area under the curve needing to be manipulated to form an estimated value for pain; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age , height, weight and baseline values for pain; Group 1 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination; Group 2 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastrointestinal adverse events (this includes but is not entirely epigastric pain, diarrhoea and nausea) at 12 weeks; Group 1: 81/337, Group 2: 14/56; Comments: DRC150 = 26.1% (29), DRC75 = 26.3% (30), EC50 = 19.6% (22), Placebo = 25.0% (14). Also reports the number of events of epigastric pain (DRC150 12.6%, DRC75 15.8%, EC50 9.8%, placebo 12.5%), diarrhoea (DRC150 6.3%, DRC75 4.4%, EC50 3.6%, placebo 0%) and nausea (DRC150 3.6%, DRC75 2.6%, EC50 5.4%, placebo 5.4%).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age , height and weight; Group 1 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination; Group 2 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination
- Actual outcome for Other: Gastric or duodenal ulcers at 12 weeks; Group 1: 2/337, Group 2: 0/56; Comments: DRC75 = 1 duodenal ulcer, EC50 = 1 gastric ulcer

Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age , height and weight; Group 1 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination; Group 2 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Other: Abnormal liver function (including elevated liver enzymes) at 12 weeks; Group 1: 26/337, Group 2: 3/56; Comments: DRC150 = 9% (10), DRC75 = 5.4% (6), EC50 = 9% (10), placebo = 5.4% (3) Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age , height and weight; Group 1 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination; Group 2 Number missing: -, Reason: 393 were randomised. Of those 89 did not complete the study (22%). 8 took excluded co-medication, 34 withdrew consent, 43 withdrew due to adverse events (17 from DRC150, 12 from DRC75, 10 from EC50, 4 from placebo), 9 dropped out for other reasons (for some there were multiple reasons for study termination	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Schnitzer 1994[169]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=59)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 9 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis in one or both hands - no additional information
Stratum	Hand
Subgroup analysis within study	Not applicable
Inclusion criteria	People, at least 18 years of age, with moderate to severe osteoarthritis hand pain
Exclusion criteria	People with underlying conditions other than arthritis requiring treatment with medication whose condition was not stable or used medications that interfered with the study medication
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 68.0 (8.8). Gender (M:F): 19:40. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hand
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 10.2 (7.6) years
Indirectness of population	No indirectness
Interventions	(n=29) Intervention 1: Capsaicin cream (topical - local) - Capsaicin cream (topical). 0.025% capsaicin cream. For 21 days applied four times a day, then for the remainder of the study for two times a day.. Duration 9 weeks. Concurrent medication/care: People taking medication for other conditions could be continue their medication. Indirectness: No indirectness (n=30) Intervention 2: Placebo. Placebo vehicle cream applied four times a day for the first 21 days of the study, then twice a day for the rest of the study. Duration 9 weeks. Concurrent medication/care: People taking medication for other conditions could be continue their medication. Indirectness: No indirectness

Funding	Study funded by industry (Supported by GenDerm Corporation, Lincolnshire, IL, and Knoll Pharmaceutical Company, Whippany, NJ)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CAPSAICIN CREAM (TOPICAL) versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Hand: Visual analogue scale at 9 weeks; Group 1: mean 33.4 (SD 14.4); n=19, Group 2: mean 37.7 (SD 23.9); n=22; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Report mean and SE values. Reported capsaicin: 33.4 (3.3). Reported vehicle: 37.7 (5.1). Baseline capsaicin: 55.0 (4.4). Baseline vehicle: 51.9 (2.8). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are not significantly different at baseline, but they are as large as the effect size and so will have an effect on the interpretation of the results.; Group 1 Number missing: 10, Reason: No additional information, was not due to adverse events; Group 2 Number missing: 8, Reason: No additional information, was not due to adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Schnitzer 2004[170]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=583)
Countries and setting	Conducted in Belgium, Denmark, Germany, Hungary, Switzerland, United Kingdom, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical or radiographic diagnosis of primary osteoarthritis of the knee or hip (diagnosed according to the American College of Rheumatology criteria)
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults (18-75 years) with clinical or radiographic diagnosis of primary osteoarthritis of the knee or hip, had been symptomatic for at least 3 months prior to enrollment, and had received an NSAID or other analgesic therapy on a regular basis. People were required to have experienced a pain intensity of at least 40mm on a 100mm visual analogue scale in the most severely affected joint during the 24 hours prior to randomisation.
Exclusion criteria	Secondary osteoarthritis or had a history of, or evidence of, specified confounding disorders (e.g. septic arthritis, inflammatory joint disease, gout, Paget's disease or articular fracture); significant medical problems that would, in the opinion of the investigator, influence outcomes; history of gastrointestinal bleeding; open knee or hip surgery within 1 year prior to study entry; anaemia; hepatic, renal or blood coagulation disorders; women who were pregnant, lactating or who were not using adequate birth control (and were of childbearing age); potentially confounding concomitant treatment after the washout period (including physiotherapy, any NSAIDs, systemic corticosteroids, intraarticular injection into the study joint, chondroitin sulfate, glucosamine sulfate, diacerhein, minocycline, histamine-2 receptor blockers, proton pump inhibitors, misoprostol, or antacids.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.3 (9.2). Gender (M:F): 188:395. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Mixed 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).

Extra comments	Severity: Not stated Duration of symptoms (mean [range]): 6.9 (0-55) years.
Indirectness of population	No indirectness
Interventions	(n=94) Intervention 1: NSAIDs - Diclofenac. Diclofenac 75mg twice a day with matching placebo. Duration 4 weeks. Concurrent medication/care: People were permitted a maximum of 6 tablets (total dose 3 grams) paracetamol per day as rescue medication.. Indirectness: No indirectness (n=97) Intervention 2: Placebo. Matching placebo twice a day. Duration 4 weeks. Concurrent medication/care: People were permitted a maximum of 6 tablets (total dose 3 grams) paracetamol per day as rescue medication.. Indirectness: No indirectness (n=390) Intervention 3: NSAIDs - Other. Valdecoxib 50mg, 100mg or 200mg twice a day or 400mg once a day. Duration 4 weeks. Concurrent medication/care: People were permitted a maximum of 6 tablets (total dose 3 grams) paracetamol per day as rescue medication.. Indirectness: No indirectness Comments: Valdecoxib is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness.
Funding	Other author(s) funded by industry (Victor S. Sloan was funded by Novartis Pharmaceuticals Corporation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 4 weeks; Group 1: mean 6.2 (SD 4); n=92, Group 2: mean 8.1 (SD 3.6); n=93; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline diclofenac: 9.5 (3.4). Baseline placebo: 9.6 (3.5).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and disease duration; Group 1 Number missing: 16, Reason: 13 adverse events, 2 protocol violations, 1 administrative problems; Group 2 Number missing: 12, Reason: 4 adverse events, 6 unsatisfactory therapeutic effect, 1 protocol violation, 1 withdrew consent

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 4 weeks; Group 1: mean 20 (SD 13.2); n=90, Group 2: mean 27.3 (SD 12.7); n=94; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline diclofenac: 31.1 (12.6). Baseline placebo: 32.7 (10.4).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover -

Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and disease duration; Group 1 Number missing: 16, Reason: 13 adverse events, 2 protocol violations, 1 administrative problems; Group 2 Number missing: 12, Reason: 4 adverse events, 6 unsatisfactory therapeutic effect, 1 protocol violation, 1 withdrew consent Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Oedema at 4 weeks; Group 1: 5/94, Group 2: 4/97 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and disease duration; Group 1 Number missing: 16, Reason: 13 adverse events, 2 protocol violations, 1 administrative problems; Group 2 Number missing: 12, Reason: 4 adverse events, 6 unsatisfactory therapeutic effect, 1 protocol violation, 1 withdrew consent	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Schnitzer 2010[173]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=918)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee, as confirmed by radiographs and the American College of Rheumatology guidelines having global functional status I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females age no less than 40 years with primary osteoarthritis of the knee. People were required to be current users of NSAIDs or paracetamol for osteoarthritis pain (i.e., had used NSAIDs or paracetamol at full therapeutic doses for at least 20 days out of 30 days during the last month prior to the screening visit) and were to have experienced a flare of pain after a discontinuation period of no less than 5 half lives of a prior analgesic or anti-inflammatory therapy.
Exclusion criteria	Uncontrolled hypertension or uncontrolled diabetes (as judged by the investigator); hepatic dysfunction or renal impairment at screening; recent coronary heart disease or stroke history within the preceding year; gastroduodenal bleeding or ulceration history within the prior 6 months; medical or arthritic disease that could interfere with efficacy evaluations; acute ligamentous or meniscal injury of the study joint within 2 years; arthroscopy of the study joint within 6 months; candidates for imminent joint replacement surgery (within 3 months)
Recruitment/selection of patients	People were required to have experienced a flare of pain after a discontinuation period of no less than 5 half lives of a prior analgesic or anti-inflammatory therapy
Age, gender and ethnicity	Age - Mean (SD): 61.4 (9.25). Gender (M:F): 274:643. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Global functional status I-III Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=227) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day. Duration 13 weeks. Concurrent medication/care: Rescue analgesia (paracetamol 500mg tablets) were provided for use in case of increased osteoarthritis pain, with a maximum accepted dose of 2000mg/day. Indirectness: No indirectness (n=221) Intervention 2: Placebo. Placebo twice a day. Duration 13 weeks. Concurrent medication/care: Rescue analgesia (paracetamol 500mg tablets) were provided for use in case of increased osteoarthritis pain, with a maximum accepted dose of 2000mg/day. Indirectness: No indirectness (n=470) Intervention 3: NSAIDs - Other. Naproxcinod 750mg twice a day or naproxcinod 375mg twice a day. Duration 13 weeks. Concurrent medication/care: Rescue analgesia (paracetamol 500mg tablets) were provided for use in case of increased osteoarthritis pain, with a maximum accepted dose of 2000mg/day. Indirectness: No indirectness Comments: Naproxcinod is not licensed for use in the UK so is not included in the analysis as agreed in the protocol. It is reported here for completeness.
Funding	Principal author funded by industry (Dr Schnitzer reports receiving research support from the National Institutes of Health, Pfizer Laborities Inc., Wyeth Laboratories, Nordic Biosciences, Novartis Pharmaceuticals Inc., Genzyme and Pozer. Dr Schnitzer presently serves as a consultant to Logical Therapeutics Inc., NicOx, Merck & Co Inc, Santosolve, Solstice and Horizon Therapeutics and is a non-invested shareholder or NicOx SA. Dr Kivitz reports receiving research support from NicOx and serves as a consultant to NicOx. Mrs Frayssinet and Dr Duquesroix are full-time employees of NicOx.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: SF-36 mental component summary score at 13 weeks; Group 1: mean 2.58 (SD 8.4); n=185, Group 2: mean 1.99 (SD 8.4); n=142; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports least square mean (standard error). Converted into standard deviation. Reported naproxen: 2.58 (0.616). Reported placebo: 1.99 (0.704). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, ACR classification for global functional status, aspirin use, diabetes, hypertension, and baseline values for WOMAC scores. Does not report baseline values for SF-36.; Group 1 Number missing: 42, Reason: 3 lost to follow up, 17 adverse events, 12 lack of efficacy/worsening, 10 others; Group 2 Number missing: 81, Reason: 7 lost to

follow up, 15 adverse events, 41 lack of efficacy/worsening, 18 other
- Actual outcome for Knee: SF-36 physical component summary score at 13 weeks; Group 1: mean 8.98 (SD 8.34); n=185, Group 2: mean 5.25 (SD 8.34); n=142; SF-36 phyiscal component summary 0-100 Top=High is good outcome; Comments: Reports least square mean (standard error). Converted into standard deviation. Reported naproxen: 8.98 (0.613). Reported placebo: 5.25 (0.700). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, ACR classification for global functional status, aspirin use, diabetes, hypertension, and baseline values for WOMAC scores. Does not report baseline values for SF-36.; Group 1 Number missing: 42, Reason: 3 lost to follow up, 17 adverse events, 12 lack of efficacy/worsening, 10 others; Group 2 Number missing: 81, Reason: 7 lost to follow up, 15 adverse events, 41 lack of efficacy/worsening, 18 other

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -36.51 (SD 27.194); n=226, Group 2: mean -24.08 (SD 27.402); n=221; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Using mean (SD). Baseline naproxen: 71.01 (17.177). Baseline placebo: 72.15 (15.831).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, ACR classification for global functional status, aspirin use, diabetes, hypertension, and baseline values for WOMAC scores. Does not report baseline values for SF-36.; Group 1 Number missing: 42, Reason: 3 lost to follow up, 17 adverse events, 12 lack of efficacy/worsening, 10 others; Group 2 Number missing: 81, Reason: 7 lost to follow up, 15 adverse events, 41 lack of efficacy/worsening, 18 other

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean -34.07 (SD 27.159); n=226, Group 2: mean -20 (SD 27.182); n=221; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Using mean (SD). Baseline naproxen: 71.05 (17.448). Baseline placebo: 70.39 (17.760).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, ACR classification for global functional status, aspirin use, diabetes, hypertension, and baseline values for WOMAC scores. Does not report baseline values for SF-36.; Group 1 Number missing: 42, Reason: 3 lost to follow up, 17 adverse events, 12 lack of efficacy/worsening, 10 others; Group 2 Number missing: 81, Reason: 7 lost to follow up, 15 adverse events, 41 lack of efficacy/worsening, 18 other

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Peripheral oedema at 13 weeks; Group 1: 9/225, Group 2: 4/222
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, ACR classification for global functional status, aspirin use, diabetes, hypertension, and baseline values for WOMAC scores. Does not report baseline values for SF-36.; Group 1 Number missing: 42, Reason: 3 lost to follow up, 17 adverse events, 12 lack of efficacy/worsening, 10 others; Group 2 Number missing: 81, Reason: 7 lost to follow up, 15 adverse events, 41 lack of efficacy/worsening, 18 other

Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Schnitzer 2011[171]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1262)
Countries and setting	Conducted in Canada, Germany, Italy, United Kingdom, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic primary osteoarthritis of the hip according to the American College of Rheumatology criteria who had experienced symptoms for at least 3 months prior to screening
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients aged ≥40 years with symptomatic primary osteoarthritis of the hip according to the American College of Rheumatology criteria and who had experienced symptoms for at least 3 months prior to screening. People were to have taken NSAIDs or simple analgesic therapy for at least 50% of the time in the previous month and were expected to require NSAID therapy for a period of at least 13 weeks. An additional requirement for randomisation was that people were to have the highest pain intenisty in the target hip joint relative to other osteoarthritis joints, have an OA pain intensity of >40mm (100mm visual analogue scale) in the target hip within 24 hours before baseline, and an increase of ≥20% and ≥10mm in OA pain intensity since screening. A similar distribution of pain severity at baseline was therefore expected for each treatment group.
Exclusion criteria	OA pain intensity of the knee(s) ≥30mm (0-100mm VAS) at screening; symptomatic osteoarthritis of the contralateral hip or spine that may interfere with assessment of the target hip; secondary osteoarthritis with a history and/or any evidence of significant diseases in the affected joints; primary fibromyalgia; rheumatoid arthritis or other inflammatory joint disease; open knee/hip surgery within the last year; observational arthroscopy, arthroscopic surgery or lavage within the last 180 days (in the knee or hip); expected to have replacement of the target joint within 4 months of enrollment; complete, even if focal, loss of articular cartilage on weight-bearing X-ray of the target joint; symptomatic trochanteric bursitis of the target joint; pregnant or lactating; peptic ulceration within the previous 12 months; clinically significant gastrointestinal bleeding within the last 5 years; evidence of hepatic, renal, blood coagulation, or anaemic disorders; a history of cardiac, cerebral thrombotic/ischaemic, or hepatic disease;

	hypersensitivity to analgesics, antipyretics, or NSAIDs; or significant medical problems; those who used rescue medication (paracetamol) >3g/day for four or more consecutive days during the screening period.
Recruitment/selection of patients	One of the inclusion criteria requires an increase of pain between screening and baseline after a washout of previous medication for up to 7 days (if people were taking piroxicam this was extended to 14 days).
Age, gender and ethnicity	Age - Mean (SD): 61.6 (10.0). Gender (M:F): 484:778. Ethnicity: White Caucasian = 1184, Black/African American = 49, Others = 29
Further population details	1. Age: Mixed (Based on range, 30-90). 2. Diagnostic method: Diagnosed with imaging (Required for the exclusion criteria). 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 3.9 (5.2) years
Indirectness of population	No indirectness
Interventions	(n=419) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 13 weeks. Concurrent medication/care: The use of rescue medication paracetamol was permitted up to a maximum of 3g/day during the study. People who took >3g of paracetamol daily for four or more continuous days during the study were to be discontinued from the trial due to "unsatisfactory therapeutic effect". No analgesic rescue medications were to be taken in 24 hours before a study visit.. Indirectness: No indirectness (n=427) Intervention 2: NSAIDs - Other. Lumiracoxib 100mg once a day. Duration 13 weeks. Concurrent medication/care: The use of rescue medication paracetamol was permitted up to a maximum of 3g/day during the study. People who took >3g of paracetamol daily for four or more continuous days during the study were to be discontinued from the trial due to "unsatisfactory therapeutic effect". No analgesic rescue medications were to be taken in 24 hours before a study visit.. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK so it was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=416) Intervention 3: Placebo. Placebo once daily. Duration 13 weeks. Concurrent medication/care: The use of rescue medication paracetamol was permitted up to a maximum of 3g/day during the study. People who took >3g of paracetamol daily for four or more continuous days during the study were to be discontinued from the trial

	due to "unsatisfactory therapeutic effect". No analgesic rescue medications were to be taken in 24 hours before a study visit.. Indirectness: No indirectness
Funding	Study funded by industry (The study was sponsored by Novartis. Individual authors received funds from other organisations as well (Genzyme, Lilly, Pfizer, Merck, NicOx).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hip: WOMAC pain subscale at 13 weeks; MD; -1.20 (95%CI -1.72 to -0.67) WOMAC pain subscale 0-20 Top=High is poor outcome, Comments: Baseline celecoxib: 10.8 (2.98). Baseline placebo: 10.6 (2.99).;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hip: WOMAC function subscale at 13 weeks; MD; -3.92 (95%CI -5.59 to -2.24) WOMAC function subscale 0-68 Top=High is poor outcome, Comments: Baseline celecoxib: 37.2 (10.45). Baseline placebo: 37.2 (10.84).;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hip: Gastrointestinal events excluding ulcers (including diarrhoea, dyspepsia, nausea) at 13 weeks; Group 1: 45/419, Group 2: 44/416; Comments: Celecoxib: 45 (including diarrhoea = 9, dyspepsia = 13, nausea = 12). Placebo: 44 (including diarrhoea = 17, dyspepsia = 11, nausea = 10).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results

- Actual outcome for Hip: Haematemesis and hematochezia at 13 weeks; Group 1: 1/419, Group 2: 2/416; Comments: Celecoxib: 1 haematemesis from a possibly nonulcerative upper GI complication. Placebo: 2 haematochezia. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Hip: Newly occurring elevations in alanine aminotransferase greater than 3 times the upper limit of normal at 13 weeks; Group 1: 0/419, Group 2: 2/416 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Hip: Headache at 13 weeks; Group 1: 47/419, Group 2: 46/416 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, region, BMI, duration of osteoarthritis and baseline values of outcomes; Group 1 Number missing: 92, Reason: 92/419 (22.0%) discontinued. 32 due to unsatisfactory therapeutic effect, 22 adverse events, 12 protocol violations, 14 withdrew consent, 6 administrative problems, 3 lost to follow-up, 1 abnormal lab value(s) or test results, 2 deaths; Group 2 Number missing: 129, Reason: 129/416 (31.0%) discontinued. 80 unsatisfactory therapeutic effect, 18 adverse events, 11 protocol violations, 11 withdrew consent, 2 administrative problems, 6 lost to follow-up, 1 abnormal lab value(s) or test results	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months

Study	Schnitzer 2011[172]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1000)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks (and an extension phase where the randomisation was undone and so will not be included in this analysis)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical diagnosis of primary osteoarthritis of the knee fulfilling the American College of Rheumatology classification criteria with American College of Rheumatology global functional status I-III and Kellgren Lawrence radiographic grade 1-3
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged ≥40 years with clinical diagnosis of primary osteoarthritis of the knee, fulfilling the American College of Rheumatology classification criteria. People with current chronic NSAID or paracetamol use for osteoarthritis pain (at least 20 days within the previous month) and had experienced a flare of pain (baseline WOMAC pain question 1 ≥50mm with an increase of ≥15mm compared with screening) after discontinuing any analgesic therapy at screening (for at least 5 half-lives of the prior analgesic or anti-inflammatory therapy before the baseline visit).
Exclusion criteria	Uncontrolled diabetes; hepatic impairment; renal impairment; clinically relevant abnormal electrocardiogram; gastroduodenal bleeding or ulceration in the previous 6 months; a history of drug abuse of addiction within 2 years of study entry; using erectile dysfunction phosphodiesterase type V inhibitors, nitrates or NO donating drugs (other than the study medication) or anticoagulants; any people who were candidates for joint replacement; if they had oral, intramuscular or lower limb intra-articular corticosteroids within 3 months; lower limb hyaluronan injections within 6 months of baseline; concomitant hypertension was allowed if it was controlled at study entry with a stable antihypertensive regimen in the 3 months before study entry; analgesics other than the study medication were not allowed during the study expect for low-dose aspirin (≤325mg/day) for cardioprotection and rescue medication (paracetamol up to 2000 mg/day); glucosamine and chondroitin were allowed only if the dose was stable for 3 months before screening/baseline and continued for the entire study duration.

Recruitment/selection of patients	Flare criteria (see inclusion criteria0
Age, gender and ethnicity	Age - Mean (SD): 59.8 (9.78). Gender (M:F): 291:709. Ethnicity: White = 789, Nonwhite = 211
Further population details	1. Age: Mixed (Based on range (40.2-89.0)). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Low comorbidity score (48% have hypertension. Does not state about other comorbidities.). 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III, Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=256) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice daily. Duration 13 weeks. Concurrent medication/care: Analgesics other than the study medication were not allowed during the study expect for low-dose aspirin (≤325mg/day) for cardioprotection and rescue medication (paracetamol up to 2000 mg/day); glucosamine and chondroitin were allowed only if the dose was stable for 3 months before screening/baseline and continued for the entire study duration.. Indirectness: No indirectness (n=489) Intervention 2: NSAIDs - Other. Naproxcinod 750mg twice a day, or Naproxcinod 375mg twice a day. Duration 13 weeks. Concurrent medication/care: Analgesics other than the study medication were not allowed during the study expect for low-dose aspirin (≤325mg/day) for cardioprotection and rescue medication (paracetamol up to 2000 mg/day); glucosamine and chondroitin were allowed only if the dose was stable for 3 months before screening/baseline and continued for the entire study duration.. Indirectness: No indirectness Comments: Naproxcinod is not licensed for use in the U and so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=257) Intervention 3: Placebo. Placebo twice a day. Duration 13 weeks. Concurrent medication/care: Analgesics other than the study medication were not allowed during the study expect for low-dose aspirin (≤325mg/day) for cardioprotection and rescue medication (paracetamol up to 2000 mg/day); glucosamine and chondroitin were allowed only if the dose was stable for 3 months before screening/baseline and continued for the entire study duration.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by NicOx)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -29.5 (SD 25.8); n=254, Group 2: mean -20.4 (SD 26); n=257; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and SE. Converted to SD. Reported naproxen: -29.5 (1.62). Reported placebo: -20.4 (1.62). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race category, BMI, American College of Rheumatology functional status and people who were hypertensive. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 65, Reason: 65 discontinued (reasons not given); Group 2 Number missing: 69, Reason: 69 discontinued (reasons not given)

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean -26.2 (SD 25); n=254, Group 2: mean -14.9 (SD 25); n=257; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square mean and SE. Reported naproxen: -26.2 (1.57). Reported placebo: -14.9 (1.56). Baseline values not reported.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race category, BMI, American College of Rheumatology functional status and people who were hypertensive. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 65, Reason: 65 discontinued (reasons not given); Group 2 Number missing: 69, Reason: 69 discontinued (reasons not given)

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: GI adverse events at 13 weeks; Group 1: 45/256, Group 2: 31/257; Comments: Naproxen: 17.5% (45). Placebo: 11.9% (31)
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race category, BMI, American College of Rheumatology functional status and people who were hypertensive. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 65, Reason: 65 discontinued (reasons not given); Group 2 Number missing: 69, Reason: 69 discontinued (reasons not given)
- Actual outcome for Knee: Severe lower GI haemorrhage at 13 weeks; Group 1: 1/256, Group 2: 0/257
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race category, BMI, American College of Rheumatology functional status and people who were hypertensive. Does not report baseline values for WOMAC subscales.; Group 1 Number missing: 65, Reason: 65 discontinued (reasons not given); Group 2 Number missing: 69, Reason: 69 discontinued (reasons not given)

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: CV adverse events at 13 weeks; Group 1: 8/256, Group 2: 11/257; Comments: Naproxen: 3.1% (8). Placebo: 4.2% (11).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, race category, BMI, American College of Rheumatology functional status and people who were hypertensive. Does not report baseline values for WOMAC subscales.; Group 1

Number missing: 65, Reason: 65 discontinued (reasons not given); Group 2 Number missing: 69, Reason: 69 discontinued (reasons not given)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Schubiger 1980[175]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=148)
Countries and setting	Conducted in Austria, Belgium, Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinically diagnosed osteoarthritis of the hip and/or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Co-operative outpatients of either sex, 18 years or over who were clinically diagnosed as suffering from osteoarthritis of the hip and/or knee
Exclusion criteria	People with concomitant conditions such as arthritis of any other aetiology (e.g. rheumatoid arthritis, gout, ankylosing spondylitis, extra-articular rheumatism, ect.), severe hepatic or renal disease; uncompensated cardiac insufficiency or severe hypertension; alcoholism with clinical manifestations; known gastric or duodenal ulcerative colitis; pregnancy; malabsorption syndrome; any abnormality in a laboratory examination severe enough in the opinion of the physician to warrant the person's exclusion
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Other: Median: 58. Gender (M:F): 47:101. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip and/or knee (some with spondyloarthrosis)).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=114) Intervention 1: NSAIDs - Diclofenac. Diclofenac 100mg sustained release (either in the morning or in the evening) or diclofenac 50mg twice a day. Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=34) Intervention 2: Placebo. Placebo. Duration 2 weeks. Concurrent

	medication/care: No additional information. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Morning pain (visual analogue scale) at 2 weeks; Group 1: mean -28 (SD 21); n=114, Group 2: mean -24 (SD 21); n=34; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports one SD for everything. Reported SD: 21. Diclofenac 100mg OM: -25, Diclofenac 100mg ON: -37, Diclofenac 50mg BD: -22. Placebo: -24.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Unclear if randomised. Reports baseline characteristics per study site.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Number of people with ≥1 GI effect at 2 weeks; Group 1: 10/114, Group 2: 3/34; Comments: Diclofenac 100mg OM: 2. Diclofenac 100mg ON: 3. Diclofenac 50mg BD: 5. Placebo: 3.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Unclear if randomised. Reports baseline characteristics per study site.; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome for Other: GI tract bleeding, perforation or ulceration at 2 weeks; Group 1: 0/114, Group 2: 0/34; Comments: "No cases of GI tract bleeding, perforation or ulceration were reported"
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Unclear if randomised. Reports baseline characteristics per study site.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Number of people with ≥1 CNS effect at 2 weeks; Group 1: 6/114, Group 2: 0/34; Comments: Diclofenac 100mg OM: 3, Diclofenac 100mg ON: 2, Diclofenac 50mg BD: 1.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Unclear if randomised. Reports baseline characteristics per study site.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-

	months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months

Study	Scott 2000[176]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=812)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: Up to 5 years (reports outcomes for 4 weeks and 12 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: All had symptomatic and radiological evidence of osteoarthritis in one or both knee joints, their clinical features were in accordance with the description of osteoarthritis in UK and North American clinical guidelines
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females with symptomatic and radiological evidence of osteoarthritis in one or both knee joints, their clinical features were in accordance with the description of osteoarthritis in UK and North American clinical guidelines
Exclusion criteria	Previous dyspeptic problems with NSAIDs; endo-stage radiological joint destruction
Recruitment/selection of patients	There was a 3-day washout period to eliminate previous anti-inflammatory and analgesic therapy; this period was longer for people receiving drugs with a long half-life, such as piroxicam. There was no specific inclusion criteria based on this.
Age, gender and ethnicity	Age - Median (range): 61 (27-87). Gender (M:F): 240:572. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms (median [range]): 5 (0.1-50) years. This study includes a follow up crossover study. This was not included in the analysis as this study only included people from the tiaprofenic acid and placebo arms, undoing the randomisation.
Indirectness of population	No indirectness
Interventions	(n=202) Intervention 1: NSAIDs - Indomethacin. Indomethacin 25mg three times a day. Duration Up to 5 years. Concurrent medication/care: Paracetamol tablets were available to all people throughout the study as escape analgesics, and people were allowed to take up to 8 tablets per day. Other drugs were used as required and their use and any change of use recorded.. Indirectness: No indirectness

	(n=303) Intervention 2: Placebo. Matching placebo. Duration Up to 5 years. Concurrent medication/care: Paracetamol tablets were available to all people throughout the study as escape analgesics, and people were allowed to take up to 8 tablets per day. Other drugs were used as required and their use and any change of use recorded.. Indirectness: No indirectness (n=307) Intervention 3: NSAIDs - Other. Tiaprofenic acid twice a day. Duration Up to 5 years. Concurrent medication/care: Paracetamol tablets were available to all people throughout the study as escape analgesics, and people were allowed to take up to 8 tablets per day. Other drugs were used as required and their use and any change of use recorded.. Indirectness: No indirectness Comments: Tiaprofenic acid is not licensed for use in the UK and so was not included in the analysis as agreed in the protocol. It is reported here for completeness.
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: INDOMETHACIN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Visual analogue scale pain score at 4 weeks; Group 1: mean 50.1 (SD 27.6); n=202, Group 2: mean 54.8 (SD 30.2); n=303; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Reports mean score and 95% confidence intervals. Standard deviation calculated from this. Reported indomethacin: 50.1 (46.3 to 53.9). Reported placebo: 54.8 (51.4 to 58.2). Baseline indomethacin: 54.9 (52.0 to 57.8). Baseline placebo: 55.0 (51.7 to 58.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Trial was terminated early due to evidence of radiographic progression in people taking indomethacin; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Does report baseline values for pain.; Group 1 Number missing: 157, Reason: 22 lack of effect, 55 adverse events, early termination in 80 people due to greater radiological progression; Group 2 Number missing: 129, Reason: 70 lack of effect, 59 adverse events

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Digestive adverse events at 20 months; Group 1: 94/202, Group 2: 98/303
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Trial was terminated early due to evidence of radiographic progression in people taking indomethacin; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Does report baseline values for pain.; Group 1 Number missing: 157, Reason: 22 lack of effect, 55 adverse events, early termination in 80 people due to greater radiological progression; Group 2 Number missing: 129, Reason: 70 lack of effect, 59 adverse events
- Actual outcome for Knee: Gastrointestinal bleeding at 20 months; Group 1: 1/202, Group 2: 0/303; Comments: Indomethacin: one episode of haematemesis
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Trial was terminated early due to evidence of radiographic progression in people taking indomethacin; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Does report baseline values for pain.; Group 1 Number missing: 157, Reason: 22 lack of effect, 55 adverse events, early termination in 80 people due to greater radiological progression; Group 2 Number missing: 129, Reason: 70 lack of effect, 59 adverse events Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiovascular adverse events at 20 months; Group 1: 12/202, Group 2: 8/303 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Trial was terminated early due to evidence of radiographic progression in people taking indomethacin; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Does report baseline values for pain.; Group 1 Number missing: 157, Reason: 22 lack of effect, 55 adverse events, early termination in 80 people due to greater radiological progression; Group 2 Number missing: 129, Reason: 70 lack of effect, 59 adverse events Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous adverse events at 20 months; Group 1: 51/202, Group 2: 51/303; Comments: Indomethacin: Nervous adverse events: 51. Placebo: Nervous adverse events: 51. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Trial was terminated early due to evidence of radiographic progression in people taking indomethacin; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only. Does report baseline values for pain.; Group 1 Number missing: 157, Reason: 22 lack of effect, 55 adverse events, early termination in 80 people due to greater radiological progression; Group 2 Number missing: 129, Reason: 70 lack of effect, 59 adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study (subsidiary papers)	Serrie 2017[177] (Biondi 2015[28], Blondi 2010[30], Etropolski 2014[65], Lange 2010[111], Lange 2018[112], Lange 2017[113])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=990)
Countries and setting	Conducted in Austria, Croatia, Germany, Hungary, Italy, Latvia, Netherlands, Poland, Portugal, Romania, Slovakia, Spain, United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 15 weeks (3 weeks of titration, 12 weeks of maintenance treatment)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a diagnosis of knee osteoarthritis based on the American College of Rheumatology criteria and functional capacity class I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People ≥40 years of age with a diagnosis of knee osteoarthritis based on the American College of Rheumatology classification criteria and functional capacity class I-III with pain requiring analgesic medications (non-opioids or opioids) at the reference joint for ≥3 months. Opioids had to be taken at doses equivalent to ≤160mg oral morphine/day. Dissatisfaction with current analgesic treatment and an average pain intensity baseline score of ≥5 on a numerical rating scale (0-10) during the last 3 days prior to randomisation.
Exclusion criteria	Clinically significant medical or psychiatric illnesses or required painful procedures (e.g. surgery) during the study that might affect efficacy or safety assessments; previously participated in a tapentadol study; participated in another clinical trial within 30 days of study enrolment; history of substance abuse; chronic hepatitis B or C; HIV infection; active hepatitis B or C within the previous 3 months; seizure disorder/epilepsy; traumatic brain injury; stroke; transient ischaemic attack; brain neoplasm; malignancy (within the previous 2 years); uncontrolled hypertension; severe renal impairment; moderate or severe hepatic impairment; ALT or AST concentrations >3 times the upper limit of normal; a clinically relevant history of hypersensitivity to the study medications or their excipients; conditions potentially influencing the assessment of osteoarthritis knee pain (e.g. anatomical deformities, significant skin conditions, fibromyalgia, gout, metabolic, infectious or autoimmune diseases at the reference joint) or surgical interventions on the reference joint within 3 month of screening or during the study; able to stop concomitant medications apart from aspirin at oral doses ≤325/day for cardiovascular prophylaxis and limited use of paracetamol as rescue medication; neuroleptics, monoamine oxidase inhibitors, serotonin norepinephrine reuptake inhibitors, tricyclic antidepressants; anti-epileptics, and anti-parkinsonian medications used within 14 days prior to screening or during the study; use of corticosteroids within 4 weeks to 6 months prior to screening, depending on the route of administration.
Recruitment/selection of patients	Recruited from trial centers across Europe (79 sites across 12 countries) from June 2007 to July 2008
Age, gender and ethnicity	Age - Mean (SD): 62.1 (9.3). Gender (M:F): 280:707. Ethnicity: Not stated

Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear (Likely diagnosed with imaging but isn't specifically confirmed). 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Majority pain severity severe, functional class I-III Duration of symptoms: Not stated explicitly. At least 3 months.. Data will be supplemented by data registered on clinicaltrials.gov (https://clinicaltrials.gov/ct2/show/study/NCT00486811?term=NCT00486811&draw=2&rank=1)
Indirectness of population	No indirectness
Interventions	(n=650) Intervention 1: Strong opioids (oral) - Tapentadol. Tapentadol PR 50mg twice daily or oxycodone CR 10mg twice daily. These were uptitrated to tapentadol PR 100mg twice daily or oxycodone CR 20mg twice daily for the next four days. These were the minimum doses for the rest of the trial but further dose increases could be achieved in 3 day intervals up to a maximum dose of tapentadol PR 250mg twice daily oxycodone CR 50mg twice daily.. Duration 3 weeks titration, 12 weeks maintenance. Concurrent medication/care: Aspirin at ≤325mg/day was allowed for cardiovascular prophylaxis. Paracetamol was allowed as rescue medications until the last 3 days of the titration period and then intermittent use for no more than 3 consecutive days was permitted during maintenance for reasons other than study-related chronic pain. Medications for psychiatric or neurological conditions not stated in the exclusion criteria were allowed if they were at a stable dose for ≥3 months prior to randomisation. Indirectness: No indirectness Comments: Both tapentadol (n=319) and oxycodone (n=331) are strong opioids and their groups have been merged as they are the same class. (n=337) Intervention 2: Placebo. Oral placebo twice daily. Duration 3 weeks titration, 12 weeks maintenance. Concurrent medication/care: Aspirin at ≤325mg/day was allowed for cardiovascular prophylaxis. Paracetamol was allowed as rescue medications until the last 3 days of the titration period and then intermittent use for no more than 3 consecutive days was permitted during maintenance for reasons other than study-related chronic pain. Medications for psychiatric or neurological conditions not stated in the exclusion criteria were allowed if they were at a stable dose for ≥3 months prior to randomisation. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Johnson & Johnson, Pharmaceutical Research & Development, L.L.C., and Global Development, Grüenthal GmbH, Aachen, Germany.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TAPENTADOL AND OXYCODONE versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D at 12 weeks; Group 1: mean 0.15 (SD 0.05); n=650, Group 2: mean 0.2 (SD 0.02); n=337; EQ-5D 0-1 Top=High is good outcome; Comments: Taken from https://clinicaltrials.gov/ct2/show/results/NCT00486811?term=NCT00486811&draw=2&rank=1. Reported least mean squares and standard deviations. Reported tapentadol: 0.2 (0.02). Reported oxycodone: 0.1 (0.01). Reported placebo: 0.2 (0.02). Baseline tapentadol: 0.4

(0.3). Baseline oxycodone: 0.4 (0.3). Baseline placebo: 0.4 (0.3).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, pain intensity, pain severity, prior opioid use, WOMAC OA global score, EQ-5D health status index, EQ-5D health state (mm), SF-36 physical component score and SF-36 mental component score. Does not report baseline values for each SF-36 subscale.; Group 1 Number missing: 343, Reason: 1 person did not receive tapentadol, 2 people did not receive oxycodone. Uses ITT for analysis. However, 343 discontinued. 201 for adverse events. 84 for patient's choice. 33 for lack of efficacy. 11 for non-compliance. 1 lost to follow up. 13 for other reasons.; Group 2 Number missing: 116, Reason: Uses all people in ITT analysis. However, 116 discontinued. 28 for adverse events. 26 as patient's choice. 43 for lack of efficacy. 5 for non-compliance. 14 for other reasons.

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain intensity (numerical rating scale) at 12 weeks; Group 1: mean -2.3 (SD 2.18); n=650, Group 2: mean -2.2 (SD 2.06); n=337; Numeric rating scale (pain) 0-10 Top=High is poor outcome; Comments: Taken from https://clinicaltrials.gov/ct2/show/results/NCT00486811?term=NCT00486811&draw=2&rank=1. Reported least mean squares and standard deviations. Reported tapentadol: -2.5 (2.18). Reported oxycodone: -2.1 (2.17). Reported placebo: -2.2 (2.06). Baseline tapentadol: 7.3 (1.1). Baseline oxcodone: 7.3 (1.1). Baseline placebo: 7.3 (1.1).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, pain intensity, pain severity, prior opioid use, WOMAC OA global score, EQ-5D health status index, EQ-5D health state (mm), SF-36 physical component score and SF-36 mental component score. Does not report baseline values for each SF-36 subscale.; Group 1 Number missing: 343, Reason: 1 person did not receive tapentadol, 2 people did not receive oxycodone. Uses ITT for analysis. However, 343 discontinued. 201 for adverse events. 84 for patient's choice. 33 for lack of efficacy. 11 for non-compliance. 1 lost to follow up. 13 for other reasons.; Group 2 Number missing: 116, Reason: Uses all people in ITT analysis. However, 116 discontinued. 28 for adverse events. 26 as patient's choice. 43 for lack of efficacy. 5 for non-compliance. 14 for other reasons.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal disorders at 12 weeks; Group 1: 414/650, Group 2: 224/337; Comments: Took the absolute numbers reported in the paper for the numbers in each group. Reported tapentadol: Absolute = 133. Serious upper abdominal pain = 1, serious constipation = 1, serious diarrhoea = 1, serious vomiting = 1, nausea = 65, constipation = 56, vomiting = 32, dry mouth = 19, diarrhoea = 14, upper abdominal pain = 11, abdominal pain = 4. Reported oxycodone: Absolute = 281. Serious constipation = 2, colonic polyp = 1, serious nausea = 1, rectal cancer = 1, nausea = 123, constipation = 114, vomiting = 86, dry mouth = 13, diarrhoea = 26, upper abdominal pain = 15, abdominal pain = 18. Reported placebo: Absolute = 224. Nausea = 21, constipation = 31, vomiting = 13, dry mouth = 7, diarrhoea = 15, upper abdominal pain = 20, abdominal pain = 7.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, pain intensity, pain severity, prior opioid use, WOMAC OA global score, EQ-5D health status index, EQ-5D health state (mm), SF-36 physical component score and SF-36 mental component score. Does not report baseline values for each SF-36 subscale.; Group 1 Number missing: 343, Reason: 1 person did not receive tapentadol, 2 people did not receive oxycodone. Uses ITT for analysis. However, 343 discontinued. 201 for adverse events. 84 for patient's choice. 33 for lack of efficacy. 11 for non-compliance. 1 lost to follow up. 13 for other reasons.; Group 2 Number missing: 116, Reason: Uses all people in ITT analysis. However, 116 discontinued. 28 for adverse events. 26 as patient's choice. 43 for lack of efficacy. 5 for non-compliance. 14 for other reasons.

Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Atrial fibrillation, myocardial infarction, paroxysmal tachycardia, ventricular arrhythmia, dyspnoea at 12 weeks; Group 1: 5/650, Group 2: 1/337; Comments: Reported tapentadol = 0. Reported oxycodone: atrial fibrillation = 3, paroxysmal tachycardia = 1, ventricular arrhythmia = 1. Reported placebo: atrial fibrillation = 1. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, pain intensity, pain severity, prior opioid use, WOMAC OA global score, EQ-5D health status index, EQ-5D health state (mm), SF-36 physical component score and SF-36 mental component score. Does not report baseline values for each SF-36 subscale.; Group 1 Number missing: 343, Reason: 1 person did not receive tapentadol, 2 people did not receive oxycodone. Uses ITT for analysis. However, 343 discontinued. 201 for adverse events. 84 for patient's choice. 33 for lack of efficacy. 11 for non-compliance. 1 lost to follow up. 13 for other reasons.; Group 2 Number missing: 116, Reason: Uses all people in ITT analysis. However, 116 discontinued. 28 for adverse events. 26 as patient's choice. 43 for lack of efficacy. 5 for non-compliance. 14 for other reasons. Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system disorders at 12 weeks; Group 1: 282/650, Group 2: 67/337; Comments: Reported tapentadol: Absolute = 130. Syncope = 1. Dizziness = 70. Somnolence = 34. Headache = 33. Reported oxycodone: Absolute = 152. Serious dizziness = 1. Dizziness = 89. Somnolence = 48. Headache = 27. Reported placebo: Absolute = 67. Dizziness = 67, somnolence = 13, headache = 31. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, pain intensity, pain severity, prior opioid use, WOMAC OA global score, EQ-5D health status index, EQ-5D health state (mm), SF-36 physical component score and SF-36 mental component score. Does not report baseline values for each SF-36 subscale.; Group 1 Number missing: 343, Reason: 1 person did not receive tapentadol, 2 people did not receive oxycodone. Uses ITT for analysis. However, 343 discontinued. 201 for adverse events. 84 for patient's choice. 33 for lack of efficacy. 11 for non-compliance. 1 lost to follow up. 13 for other reasons.; Group 2 Number missing: 116, Reason: Uses all people in ITT analysis. However, 116 discontinued. 28 for adverse events. 26 as patient's choice. 43 for lack of efficacy. 5 for non-compliance. 14 for other reasons.	
Protocol outcomes not reported by the study	Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study (subsidiary papers)	Sheldon 2005[178] (Sheldon 2008[179])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1551)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee based on American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients aged ≥18 years with a clinical diagnosis of primary osteoarthritis of the knee based on the American College of Rheumatology criteria. People were required to have symptoms necessitating NSAID treatment for ≥3 months before study entry.
Exclusion criteria	Secondary osteoarthritis; recent surgery in the target knee; open knee surgery within the past year; observational arthroscopy, arthroscopic surgery or lavage within the past 180 days; known hypersensitivity to analgesics, antipyretics or NSAIDs; other connective tissue disease; significant medical problems such as recent peptic ulceration, GI bleeding, or a history of malignancy within the past 5 years; women of childbearing potential not using a reliable method of contraception
Recruitment/selection of patients	The screening period was followed by a washout of medication for 3-7 days (the number of days determined by specific NSAID therapy at study entry) but was more than or equal to 14 days in people who been receiving piroxicam and similar agents with long half-lives (>60 hours). After the washout period, people who had pain intensity in the target knee of more than or equal to 40 mm on a 100mm visual analogue scale during the 24 hours before assessment (people were asked to describe the worst pain experienced during this period).
Age, gender and ethnicity	Age - Mean (SD): 60.5 (10.8). Gender (M:F): 583:968. Ethnicity: White = 1387, Black = 83, Hispanic = 53, Other = 28
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of osteoarthritis (mean [SD]): 7.0 (7.8) years

Indirectness of population	No indirectness
Interventions	(n=393) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 13 weeks. Concurrent medication/care: Concomitant treatment with NSAIDs or other analgesics were not allowed during the study, with the exception of low-dose aspirin (≤325mg/day) for cardiovascular prophylaxis in people at increased cardiovascular risk. Use of paracetamol (500mg tablet) as rescue medication was permitted throughout the study, including during the screening period. People consuming paracetamol ≥2g/d during the screening period were not eligible for randomisation. People consuming paracetamol ≥2g/d after the study had begun were discontinued for unsatisfactory therapeutic effect. Other concomitant drug therapies were allowed (including during the screening period) including chondroitin sulfate and/or glucosamine sulfate (if the dose and regimen were established and stable), corticosteroids (topical, ophthalamic, nasal, or inhaled, at usual labeled doses), histamine-2 receptor antagonists, proton pump inhibitors, antacids and cytoprotective agents (taken at the usual labeled doses) and physiotherapy as prescribed by the physician. People were not permitted to take diacerein or minocycline within 1 month of screening, drugs known to be contraindicated with celecoxib, systemic corticosteroids within the past month, anticoagulants, hyaluronic acid injection or intra-articular corticosteroid in the study joint in the past 3 months, and opiates taken more than twice a week during the month before screening.. Indirectness: No indirectness (n=776) Intervention 2: NSAIDs - Other. Lumiracoxib 100mg or 100mg with a loading dose. Duration 13 weeks. Concurrent medication/care: (if the dose and regimen were established and stable), corticosteroids (topical, ophthalamic, nasal, or inhaled, at usual labeled doses), histamine-2 receptor antagonists, proton pump inhibitors, antacids and cytoprotective agents (taken at the usual labeled doses) and physiotherapy as prescribed by the physician. People were not permitted to take diacerein or minocycline within 1 month of screening, drugs known to be contraindicated with celecoxib, systemic corticosteroids within the past month, anticoagulants, hyaluronic acid injection or intra-articular corticosteroid in the study joint in the past 3 months, and opiates taken more than twice a week during the month before screening.. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK and was not included in the analysis as agreed in the protocol. It was reported here for completeness. (n=382) Intervention 3: Placebo. Matching placebo once a day. Duration 13 weeks. Concurrent medication/care: (if the dose and regimen were established and stable), corticosteroids (topical, ophthalamic, nasal, or inhaled, at usual labeled doses),

	histamine-2 receptor antagonists, proton pump inhibitors, antacids and cytoprotective agents (taken at the usual labeled doses) and physiotherapy as prescribed by the physician. People were not permitted to take diacerein or minocycline within 1 month of screening, drugs known to be contraindicated with celecoxib, systemic corticosteroids within the past month, anticoagulants, hyaluronic acid injection or intra-articular corticosteroid in the study joint in the past 3 months, and opiates taken more than twice a week during the month before screening.. Indirectness: No indirectness
Funding	Principal author funded by industry (Victor S. Sloan is an employee of Novartis Phaemceuticals Corporation, East Hanover, New Jersey)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -3.4 (SD 4.21); n=393, Group 2: mean -2.3 (SD 3.84); n=382; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 10.8 (3.24). Baseline placebo: 11.0 (2.89).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, duration of osteoarthritis, pain, patient and physician global assessment of disease, WOMAC subscales and SF-36 component scores; Group 1 Number missing: 80, Reason: 80 discontinued. 16 adverse events, 32 unsatisfactory therapeutic effect, 12 protocol violation, 7 withdrawal of consent, 3 lost to follow up; Group 2 Number missing: 132, Reason: 24 adverse events, 70 unsatisfactory therapeutic effect, 19 protocol violation, 10 withdrawal of consent, 2 abnormal laboratory value(s), 6 lost to follow up, 1 administrative problem

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC difficulty in performing daily activities subscale at 13 weeks; Group 1: mean -10.8 (SD 13.07); n=393, Group 2: mean -6.3 (SD 11.8); n=382; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline celecoxib: 36.9 (10.95). Baseline placebo: 37.2 (10.47).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, duration of osteoarthritis, pain, patient and physician global assessment of disease, WOMAC subscales and SF-36 component scores; Group 1 Number missing: 80, Reason: 80 discontinued. 16 adverse events, 32 unsatisfactory therapeutic effect, 12 protocol violation, 7 withdrawal of consent, 3 lost to follow up; Group 2 Number missing: 132, Reason: 24 adverse events, 70 unsatisfactory therapeutic effect, 19 protocol violation, 10 withdrawal of consent, 2 abnormal laboratory value(s), 6 lost to follow up, 1 administrative problem

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Diarrhoea at 13 weeks; Group 1: 15/393, Group 2: 13/382
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, duration of osteoarthritis, pain, patient and physician global assessment of disease, WOMAC subscales and SF-36 component scores; Group 1 Number

missing: 80, Reason: 80 discontinued. 16 adverse events, 32 unsatisfactory therapeutic effect, 12 protocol violation, 7 withdrawal of consent, 3 lost to follow up; Group 2 Number missing: 132, Reason: 24 adverse events, 70 unsatisfactory therapeutic effect, 19 protocol violation, 10 withdrawal of consent, 2 abnormal laboratory value(s), 6 lost to follow up, 1 administrative problem Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Alanine aminotransferase/aspartate aminotransferase levels >3 times the upper limit of normal at 13 weeks; Group 1: 1/393, Group 2: 0/382 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, duration of osteoarthritis, pain, patient and physician global assessment of disease, WOMAC subscales and SF-36 component scores; Group 1 Number missing: 80, Reason: 80 discontinued. 16 adverse events, 32 unsatisfactory therapeutic effect, 12 protocol violation, 7 withdrawal of consent, 3 lost to follow up; Group 2 Number missing: 132, Reason: 24 adverse events, 70 unsatisfactory therapeutic effect, 19 protocol violation, 10 withdrawal of consent, 2 abnormal laboratory value(s), 6 lost to follow up, 1 administrative problem Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 13 weeks; Group 1: 47/393, Group 2: 46/382 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, body mass index, duration of osteoarthritis, pain, patient and physician global assessment of disease, WOMAC subscales and SF-36 component scores; Group 1 Number missing: 80, Reason: 80 discontinued. 16 adverse events, 32 unsatisfactory therapeutic effect, 12 protocol violation, 7 withdrawal of consent, 3 lost to follow up; Group 2 Number missing: 132, Reason: 24 adverse events, 70 unsatisfactory therapeutic effect, 19 protocol violation, 10 withdrawal of consent, 2 abnormal laboratory value(s), 6 lost to follow up, 1 administrative problem	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months

Study	Sikes 2002[180]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1052)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A documented clinical diagnosis of osteoarthritis
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People of the legal age for consent with a documented clinical diagnosis of osteoarthritis who required chronic use of NSAIDs and/or oral analgesics.
Exclusion criteria	More than ten gastric or duodenal erosions or if there was evidence of oesophageal, gastric, pyloric channel or duodenal ulcer during the pretreatment uppergastrointestinal endoscopy; females of childbearing age if they were pregnant, lactating or using inadequate methods of contraception; people diagnosed with any other inflammatory arthritis, active gout, active gastrointestinal disease, or active malignancy or history of malignancy; people who had taken any NSAIDs or analgesics within 48 hours of the pretreatment endoscopy or warfarin within 30 days of receiving the study medication; the use of anti-ulcer drugs, antibiotics as treatment for Helicobacter pylori infection, anticoagulants and lithium; uncontrolled hypertension; diabetes; chronic or acute renal or hepatic disorders; significant coagulantion defect; AST, ALT or blood urea nitrogen values exceeded 1.5 times the upper limit of normal, or if serum creatinine levels exceeded 1.5mg/dL.
Recruitment/selection of patients	Multicentre (72 sites in the USA, 8 sites in Canada)
Age, gender and ethnicity	Age - Other: Mean: 59.9. Gender (M:F): 335:717. Ethnicity: White = 852, Black = 108, Other = 92
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Unclear).
Extra comments	Severity: Not stated Duration of symptoms (mean): 10.4 years
Indirectness of population	No indirectness

Interventions	(n=419) Intervention 1: NSAIDs - Diclofenac. Ibuprofen 800mg three times a day or diclofenac sodium 75mg twice a day with matching placebo (double-dummy technique). Duration 12 weeks. Concurrent medication/care: Aspirin ≤325mg/day was permitted for non-arthritic reasons if being taken at a stable dose for at least 30 days before the first dose of study drug. Indirectness: No indirectness Comments: Ibuprofen and Diclofenac groups were separate but were combined together in this analysis due to class effect as agreed in the protocol. (n=423) Intervention 2: NSAIDs - Other. Valdecoxib 10mg daily and Valdecoxib 20mg daily with matching placebo (double-dummy technique). Duration 12 weeks. Concurrent medication/care: Aspirin ≤325mg/day was permitted for non-arthritic reasons if being taken at a stable dose for at least 30 days before the first dose of study drug. Indirectness: No indirectness Comments: Valdecoxib is not licensed for use in the UK so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=210) Intervention 3: Placebo. Matching placebo three times per day. Duration 12 weeks. Concurrent medication/care: Aspirin ≤325mg/day was permitted for non-arthritic reasons if being taken at a stable dose for at least 30 days before the first dose of study drug. Indirectness: No indirectness
Funding	Principal author funded by industry (Kenneth M. Varburg is an employee of Pharmacia. William W. Zhao, Jeffrey D. Kent and David P. Recker are employees of the Pharmacia Corporation. David H. Sikes and Naurang M. Agrawal have acted as consultants for Pharmacia. Therefore, all authors have received benefits or will receive benefits from this organisation.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC AND IBUPROFEN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastroduodenal ulcers at 12 weeks; Group 1: 50/419, Group 2: 8/210; Comments: Diclofenac: 25, Ibuprofen: 25, Placebo: 8
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, mean weight, Helicobacter pylori status, mean duration of osteoarthritis and history of gastroduodenal ulcers; Group 1 Number missing: 111, Reason: Diclofenac: 1 lost to follow up, 4 protocol violations, 9 non-compliance, 12 treatment failures, 34 adverse events. Ibuprofen: 1 lost to follow up, 2 protocol violations, 10 non-compliance, 11 treatment failure, 27 adverse events.; Group 2 Number missing: 75, Reason: 2 Lost to follow up, 6 protocol violations, 7 non-compliance, 45 treatment failure, 15 adverse events

Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Hypertension at 12 weeks; Group 1: 7/419, Group 2: 1/210; Comments: Ibuprofen: 5, Diclofenac: 2, Placebo: 1 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, mean weight, Helicobacter pylori status, mean duration of osteoarthritis and history of gastroduodenal ulcers; Group 1 Number missing: 111, Reason: Diclofenac: 1 lost to follow up, 4 protocol violations, 9 non-compliance, 12 treatment failures, 34 adverse events. Ibuprofen: 1 lost to follow up, 2 protocol violations, 10 non-compliance, 11 treatment failure, 27 adverse events.; Group 2 Number missing: 75, Reason: 2 Lost to follow up, 6 protocol violations, 7 non-compliance, 45 treatment failure, 15 adverse events Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 12 weeks; Group 1: 9/419, Group 2: 9/210; Comments: Diclofenac: 5, Ibuprofen: 4, Placebo: 9 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, mean weight, Helicobacter pylori status, mean duration of osteoarthritis and history of gastroduodenal ulcers; Group 1 Number missing: 111, Reason: Diclofenac: 1 lost to follow up, 4 protocol violations, 9 non-compliance, 12 treatment failures, 34 adverse events. Ibuprofen: 1 lost to follow up, 2 protocol violations, 10 non-compliance, 11 treatment failure, 27 adverse events.; Group 2 Number missing: 75, Reason: 2 Lost to follow up, 6 protocol violations, 7 non-compliance, 45 treatment failure, 15 adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study (subsidiary papers)	Simon 2009[181] (Roth 2011[156])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=775)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee based on standard radiological criteria of osteoarthritis, on a recent (within 3 months) examination pain with regular use of a NSAID or other analgesic medication (at least 3 days a week in the previous month) and a flare of pain and minimum Likert pain score of 8 (40 on a scale normalised to 0-100) at baseline following a washout of that medication.
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and non-pregnant women aged 40-85 years with primary osteoarthritis of the knee based on the diagnostic criteria (see above).
Exclusion criteria	Exclusion criteria from another study. Secondary arthritis related to systemic inflammatory arthritis, including rheumatoid arthritis, psoriatic arthritis, postinfectious arthritis, metabolic arthritis, and traumatic arthritis or surgical joint replacements; sensitivity to diclofenac, aspirin or any other NSAID, dimethyl sulfoxide, propylene glycol, glycerin, or ethanol; clinically active renal, hepatic, or peptic ulcer disease; a history of alcohol or other drug abuse; lactation; concomitant skin disease at the application site; corticosteroid use, including oral corticosteroid within 14 days, intramuscular corticosteroid within 30 days, intra-articular corticosteroid into the study knee within 90 days, intra-articular corticosteroid into any other joint within 30 days of study entry or ongoing use of topical corticosteroid at the site of application; use of a topical product, treatment, or device at the application site for the relief of OA; ongoing use of prohibited medication, including NSAIDs, oral analgesic, muscle relaxant, or low-dose antidepressant; ongoing use of glucosamine or chondroitin sulfate sodium (unless used continuously for 90 days before study entry); intra-articular viscosupplementation (e.g. hyaluronate sodium derivative) into the study knee in the preceding 90 days; current application for disability benefits on the basis of osteoarthritis of the knee; fibromyalgia; other painful or disabling condition affecting the knee

Recruitment/selection of patients	Required people to have a flare (a flare was defined as an increase in total Likert pain score of 25% and at least 2, and a score of at least moderate on one or more of the five items/questions of the WOMAC LK3.1 pain subscale).
Age, gender and ethnicity	Age - Mean (SD): 61.6 (9.9). Gender (M:F): 292:480. Ethnicity: Caucasian = 598, Black = 41, Asian = 70, Hispanic = 45, Other = 18
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Mean study knee radiographic score (mean [SD]): 16.5 (3.1) Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=154) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Topical diclofenac solution (1.5% w/w diclofenac sodium in a vehicle solution containing 45.5% w/w DMSO and other excipients) applied four times daily to the entire circumference of the study knee without massage and 1 placebo tablet daily. Duration 12 weeks. Concurrent medication/care: Concomitant analgesic and anti-inflammatory medications, including over-the-counter NSAIDs and other analgesics were prohibited. Continuation of stable treatment with glucosamine, chondroitin, anti-depressants or a proton pump inhibitor (previous 90 days), or low dose (≤325mg/day) acetylsalicylic acid (previous 30 days); paracetamol was provided and permitted (up to four 325mg caplets per day) except during the 3 days before each efficacy assessment; other topical products on the knee, including skin emollients, were prohibited. A person with a gastrointestinal adverse event was allowed to start a proton pump inhibitor. (n=318) Intervention 2: Placebo. DMSO only vehicle applied four times a day or placebo gel (without DMSO) applied four times a day both with placebo tablets once a day. Duration 12 weeks. Concurrent medication/care: Concomitant analgesic and anti-inflammatory medications, including over-the-counter NSAIDs and other analgesics were prohibited. Continuation of stable treatment with glucosamine, chondroitin, anti-depressants or a proton pump inhibitor (previous 90 days), or low dose (≤325mg/day) acetylsalicylic acid (previous 30 days); paracetamol was provided and permitted (up to four 325mg caplets per day) except during the 3 days before each efficacy assessment; other topical products on the knee, including skin emollients, were prohibited. A person with a gastrointestinal adverse event was allowed to start a proton pump inhibitor.. Indirectness: No indirectness Comments: The two groups were combined as both were considered to be placebo (n=151) Intervention 3: NSAIDs - Diclofenac. Oral diclofenac (100mg slow release)

	once per day with placebo gel four times a day. Duration 12 weeks. Concurrent medication/care: Concomitant analgesic and anti-inflammatory medications, including over-the-counter NSAIDs and other analgesics were prohibited. Continuation of stable treatment with glucosamine, chondroitin, anti-depressants or a proton pump inhibitor (previous 90 days), or low dose (≤325mg/day) acetylsalicylic acid (previous 30 days); paracetamol was provided and permitted (up to four 325mg caplets per day) except during the 3 days before each efficacy assessment; other topical products on the knee, including skin emollients, were prohibited. A person with a gastrointestinal adverse event was allowed to start a proton pump inhibitor.. Indirectness: No indirectness (n=152) Intervention 4: NSAIDs - Diclofenac. Oral diclofenac (100mg slow release) once per day with Topical diclofenac solution (1.5% w/w diclofenac sodium in a vehicle solution containing 45.5% w/w DMSO and other excipients) applied four times daily to the entire circumference of the study knee without massage. Duration 12 weeks. Concurrent medication/care: Concomitant analgesic and anti-inflammatory medications, including over-the-counter NSAIDs and other analgesics were prohibited. Continuation of stable treatment with glucosamine, chondroitin, anti-depressants or a proton pump inhibitor (previous 90 days), or low dose (≤325mg/day) acetylsalicylic acid (previous 30 days); paracetamol was provided and permitted (up to four 325mg caplets per day) except during the 3 days before each efficacy assessment; other topical products on the knee, including skin emollients, were prohibited. A person with a gastrointestinal adverse event was allowed to start a proton pump inhibitor.. Indirectness: No indirectness Comments: This group is not included in the protocol and so was not included for the full analysis. It was reported here for completeness.
Funding	Study funded by industry (The sponsor, Nuvo Research Inc., provided medication and funding for this study, and managed and analysed the data)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -6 (SD 4.5); n=154, Group 2: mean -4.7 (SD 4.5); n=318; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Diclofenac gel: -6.0 (4.5). Placebo: -4.7 (4.4). DMSO: -4.7 (4.3). Baseline diclofenac gel: 13.2 (3.4). Baseline placebo: 12.9 (3.3). Baseline DMSO: 13.0 (3.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of

hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 102, Reason: Placebo: Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -15.8 (SD 15.1); n=154, Group 2: mean -12.2 (SD 14.7); n=318; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Topical diclofenac: -15.8 (15.1). Placebo: -12.3 (14.7). DMSO: -12.1 (14.6). Baseline topical diclofenac: 41.7 (12.8). Baseline placebo: 41.6 (11.7). Baseline DMSO: 41.4 (11.4).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 102, Reason: Placebo: Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Rectal haemorrhage at 12 weeks; Group 1: 1/154, Group 2: 0/318; Comments: Diclofenac gel: 1, placebo = 0, DMSO = 0
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 102, Reason: Placebo: Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus DICLOFENAC

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -6 (SD 4.5); n=154, Group 2: mean -6.4 (SD 4.1); n=151; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Diclofenac gel: -6.0 (4.5). Oral diclofenac: -6.4 (4.1). Baseline diclofenac gel: 13.2 (3.4). Baseline oral diclofenac: 13.2 (3.0).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.

Protocol outcome 2: Physical function at ≤3- or >3- months

- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -15.8 (SD 15.1); n=154, Group 2: mean -12.3 (SD 14.7); n=151; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Topical diclofenac: -15.8 (15.1). Oral diclofenac: -12.3 (14.7). Baseline topical diclofenac: 41.7 (12.8). Baseline oral diclofenac: 42.1 (12.0).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Rectal haemorrhage at 12 weeks; Group 1: 1/154, Group 2: 0/151; Comments: Diclofenac gel: 1, diclofenac: 0
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 51, Reason: Diclofenac gel: Adverse events = 16, lack of effect = 16, consent withdrawn = 6, lost to follow up = 2, other = 11.; Group 2 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -6.4 (SD 4.1); n=151, Group 2: mean -4.7 (SD 4.4); n=318; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Oral diclofenac: -6.4 (4.1). Placebo: -4.7 (4.4). DMSO: -4.7 (4.3). Baseline oral diclofenac: 13.2 (3.0). Baseline placebo: 12.9 (3.3). Baseline DMSO: 13.0 (3.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.; Group 2 Number missing: 102, Reason: Placebo: Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -12.3 (SD 14.7); n=151, Group 2: mean -12.2 (SD 14.7); n=318; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Oral diclofenac: -12.3 (14.7). Placebo: -12.3 (14.7). DMSO: -12.1 (14.6). Baseline oral diclofenac: 42.1 (12.0). Baseline placebo: 41.6 (11.7). Baseline DMSO: 41.4 (11.4).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.; Group 2 Number missing: 102, Reason: Placebo:

Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6 Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Rectal haemorrhage at 12 weeks; Group 1: 0/151, Group 2: 0/318; Comments: diclofenac: 0, placebo = 0, DMSO = 0 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, incidence of hypertension, total study knee radiographic score, number of knees with osteoarthritis and baseline values of outcomes; Group 1 Number missing: 44, Reason: Diclofenac: Adverse events = 19, lack of effect = 5, consent withdrawn = 8, lost to follow up = 2, other = 10.; Group 2 Number missing: 102, Reason: Placebo: Adverse events = 18, lack of effect = 18, consent withdrawn = 6, lost to follow up = 4, other = 8. DMSO: Adverse events = 12, lack of effect = 17, consent withdrawn = 10, lost to follow up = 3, other = 6	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Smugar 2006[182]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	2 (2 studies pooled for analysis) (n=2594)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with clinical diagnosis of osteoarthritis of the knee or hip
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females, at least 40 years of age, with a clinical diagnosis of osteoarthritis of the knee or hip for at least 6 months and American College of Rheumatology functional Class I, II or III, requiring regular treatment with prescription-strength NSAIDs (including COX-2 inhibitors). People had to be on NSAID therapy at therapeutic doses for at least 30 days prior to enrollment. People required a pain on walking on a flat surface score (WOMAC) at the screening visit to be less than 80mm on a 100mm visual analogue scale.
Exclusion criteria	Concurrent medical (or arthritic) disease that potentially confounded the evaluation of efficacy, including inflammatory arthritis and metabolic disorders associated with arthritis; isolated patello-femoral disease manifested by primary anterior knee pain; candidacy for immenent joint replacement; history of acute injury of the study joint within the previous 2 years; clinical of laboratory evidence of systemic disease that would contradict the use of NSAIDs or paracetamol; serum creatinine greater than 2.0mg/dL; abnormal haemoglobin or haematocrit levels; angina or congestive heart failure with symptoms at rest; stroke, transient ischaemic attack, myocardial infarction or coronary artery bypass grafting within the previous year; uncontrolled hypertension; history of alcohol or substance abuse within the previous 5 years; history of allergic reaction to NSIADs, paracetamol or sulfa drugs
Recruitment/selection of patients	After the screening visit and an NSAID washout period, with a duration beyond five plasma half-lives for the NSAID each person was taking, people returned for evaluation of worsening pain or other symptoms of osteoarthritis. Based on the WOMAC questionnaire, people were eligible for randomisation if they had a minimum of 40mm on a flat surface, a minimum of 40mm on patient assessed pain at night VAS, an increase of 15mm from baseline (on therapy) on patient assessed pain on walking on a flat surface VAS, and a worsening in Investigator Global Assessment of

	Disease Status of at least one point on a five point Likert scale compared with the screening visit
Age, gender and ethnicity	Age - Mean (range): 61.7 (39-92). Gender (M:F): 859:1744. Ethnicity: Asian = 12, Black = 178, Hispanic-American = 55, Native American = 6, White = 2339, Other = 13
Further population details	1. Age: Mixed 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American College of Rheumatology function class I-III Duration of symptoms: From 6 months to >10 years. Mean 1-10 years.
Indirectness of population	No indirectness
Interventions	(n=916) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 6 weeks. Concurrent medication/care: Use of any of the following agents was not permitted: intraarticular steroids, sodium hyaluroante, or hylan G-F 20 (all if within the 3 months prior to enrollment), warfarin, ticlopidine, clopidogrel bisulfate, muscle relaxants, corticosteroids (within the previous month, excluding topical), high-dose aspirin (>81 mg/day), appetite suppressants, misoprostol or sucralfate within the past 14 days, histamine-2 receptor blockers (on a regular basis), proton pump inhibitors, calcium-containing antacids, and topical, oral or systemic analgesics other than study medications or paracetamol. Indirectness: No indirectness (n=1389) Intervention 2: NSAIDs - Other. Rofecoxib 12.5mg or 25mg once a day. Duration 6 weeks. Concurrent medication/care: Use of any of the following agents was not permitted: intraarticular steroids, sodium hyaluroante, or hylan G-F 20 (all if within the 3 months prior to enrollment), warfarin, ticlopidine, clopidogrel bisulfate, muscle relaxants, corticosteroids (within the previous month, excluding topical), high-dose aspirin (>81 mg/day), appetite suppressants, misoprostol or sucralfate within the past 14 days, histamine-2 receptor blockers (on a regular basis), proton pump inhibitors, calcium-containing antacids, and topical, oral or systemic analgesics other than study medications or paracetamol. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK so was not included in the analysis as agreed in the protocol. It is reported here for completeness. (n=301) Intervention 3: Placebo. Matching placebo once a day. Duration 6 weeks. Concurrent medication/care: Use of any of the following agents was not permitted: intraarticular steroids, sodium hyaluroante, or hylan G-F 20 (all if within the 3 months prior to enrollment), warfarin, ticlopidine, clopidogrel bisulfate, muscle relaxants, corticosteroids (within the previous month, excluding topical), high-dose aspirin (>81

	mg/day), appetite suppressants, misoprostol or sucralfate within the past 14 days, histamine-2 receptor blockers (on a regular basis), proton pump inhibitors, calcium-containing antacids, and topical, oral or systemic analgesics other than study medications or paracetamol. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Merck and Company, Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 6 weeks; Group 1: mean -31.9 (SD 21.5); n=916, Group 2: mean -19.3 (SD 22.3); n=301; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported celecoxib study 1: -33.0 (1.0). Reported celecoxib study 2: -30.8 (1.0). Reported placebo study 1: -22.0 (1.8). Reported placebo study 2: -16.7 (1.8). Calculated celecoxib SD study 1: 21.4. Calculated celecoxib SD study 2: 21.5. Calculated placebo SD study 1: 22.1. Calculated placebo SD study 2: 22.1. Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, duration of osteoarthritis, mean pain at baseline and baseline IGADS. Does not report baseline WOMAC values.; Group 1 Number missing: 148, Reason: Study 1: 80 withdrew. 45 lack of efficacy, 23 adverse events, 4 deviation from protocol, 2 lost to follow up, 3 withdrew consent, 3 other. Study 2: 68 withdrew. 42 lack of efficacy, 14 adverse events, 4 deviation from protocol, 3 lost to follow up, 4 withdrew consent, 1 other; Group 2 Number missing: 115, Reason: Study 1: 57 withdrew. 39 lack of efficacy, 9 adverse events, 2 deviation from protocol, 1 lost to follow up, 5 withdrew consent, 1 other. Study 2: 58 withdrew. 47 lack of efficacy, 3 adverse events, 2 deviation from protocol, 2 lost to follow up, 4 withdrew consent.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 6 weeks; Group 1: mean -26.4 (SD 21.5); n=916, Group 2: mean -14 (SD 22.3); n=301; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports change scores and standard error. Calculated SD from this. Reported celecoxib study 1: -27.4 (1.0). Reported celecoxib study 2: -25.3 (1.0). Reported placebo study 1: -16.4 (1.9). Reported placebo study 2: -11.6 (1.7). Calculated celecoxib SD study 1: 21.4. Calculated celecoxib SD study 2: 21.5. Calculated placebo SD study 1: 23.3. Calculated placebo SD study 2: 20.9. Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, duration of osteoarthritis, mean pain at baseline and baseline IGADS. Does not report baseline WOMAC values.; Group 1 Number missing: 148, Reason: Study 1: 80 withdrew. 45 lack of efficacy, 23 adverse events, 4 deviation from protocol, 2 lost to follow up, 3 withdrew consent, 3 other. Study 2: 68 withdrew. 42 lack of efficacy, 14 adverse events, 4 deviation from protocol, 3 lost to follow up, 4 withdrew consent, 1 other; Group 2 Number missing: 115, Reason: Study 1: 57 withdrew. 39 lack of efficacy, 9 adverse events, 2 deviation from protocol, 1 lost to follow up, 5 withdrew consent, 1 other. Study 2: 58 withdrew. 47 lack of efficacy, 3 adverse events, 2 deviation from protocol, 2 lost to follow up, 4 withdrew consent.

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Other: Gastrointestinal adverse events at 6 weeks; Group 1: 101/916, Group 2: 27/301 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, duration of osteoarthritis, mean pain at baseline and baseline IGADS. Does not report baseline WOMAC values.; Group 1 Number missing: 148, Reason: Study 1: 80 withdrew. 45 lack of efficacy, 23 adverse events, 4 deviation from protocol, 2 lost to follow up, 3 withdrew consent, 3 other. Study 2: 68 withdrew. 42 lack of efficacy, 14 adverse events, 4 deviation from protocol, 3 lost to follow up, 4 withdrew consent, 1 other; Group 2 Number missing: 115, Reason: Study 1: 57 withdrew. 39 lack of efficacy, 9 adverse events, 2 deviation from protocol, 1 lost to follow up, 5 withdrew consent, 1 other. Study 2: 58 withdrew. 47 lack of efficacy, 3 adverse events, 2 deviation from protocol, 2 lost to follow up, 4 withdrew consent. Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Discontinuations due to oedema related or hypertension related events at 6 weeks; Group 1: 5/916, Group 2: 0/301; Comments: Celecoxib study 1: Oedema-related events = 3, hypertension-related events = 1. Celecoxib study 2: Oedema-related events = 1, hypertension-related events = 0. Placebo = 0. Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Discontinuation events; Baseline details: Reports sex, age, race, duration of osteoarthritis, mean pain at baseline and baseline IGADS. Does not report baseline WOMAC values.; Group 1 Number missing: 148, Reason: Study 1: 80 withdrew. 45 lack of efficacy, 23 adverse events, 4 deviation from protocol, 2 lost to follow up, 3 withdrew consent, 3 other. Study 2: 68 withdrew. 42 lack of efficacy, 14 adverse events, 4 deviation from protocol, 3 lost to follow up, 4 withdrew consent, 1 other; Group 2 Number missing: 115, Reason: Study 1: 57 withdrew. 39 lack of efficacy, 9 adverse events, 2 deviation from protocol, 1 lost to follow up, 5 withdrew consent, 1 other. Study 2: 58 withdrew. 47 lack of efficacy, 3 adverse events, 2 deviation from protocol, 2 lost to follow up, 4 withdrew consent.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Sofat 2017[183]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=65)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis diagnosed by American College of Rheumatology criteria confirmed by a rheumatologist and experiencing pain of at least 5 on a Numerical Rating Scale of 0-10.
Stratum	Hand
Subgroup analysis within study	Not applicable
Inclusion criteria	People fulfilling the American College of Rheumatologists criteria for the diagnosis of hand osteoarthritis, male or female, right or left handed, aged 40-75 years, and on usual care for hand osteoarthritis including paracetamol and/or NSAIDs.
Exclusion criteria	Another rheumatological disease (e.g. rheumatoid arthritis); current or planned pregnancy; contraindications to duloxetine or pregabalin such as concomitant use of monoamine oxidase inhibitors, selective serotonin reuptake inhibitors, antidepressants, oral contraceptives, St. John's wort; history of depression; concomitant use of opioids including tramadol and pethidine; use of benzodiazepines; recent surgery (ie. <6 weeks prior to participation in the study); recent insertion of surgical implants (ie, <6 weeks before participation prior to entry); previous use of duloxetine and/or pregabalin; uncontrolled depression; estimated glomerular filtration rate <60mL/min; hepatic impairment defined as ALT >2.5x upper limit of normal within 6 weeks of last clinical assessment; ischaemic heart disease; diabetes mellitus; regular use of alcohol or alcohol abuse (maximum limits are 28 units/week for men and 21 units/week for women); lactose intolerance; people with a Hospital Anxiety and Depression Scale score of no less than 12 for anxiety and/or depression.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.9 (7.2). Gender (M:F): 13:52. Ethnicity: White = 58, Black = 4, Asian = 3
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Hand

Extra comments	Severity: Not stated Duration of symptoms: Not stated. NCT02612233
Indirectness of population	No indirectness
Interventions	(n=22) Intervention 1: Anti-epileptics (oral) - Pregabalin. Pregabalin 150mg tablets. 1 capsule at night for 1 week, then 2 capsules daily for 10 weeks, then 1 capsule at night for 2 weeks. Duration 13 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=21) Intervention 2: Antidepressants (oral) - SNRIs. Duloxetine 30mg tablets. 1 capsule at night for 1 week, then 2 capsules daily for 10 weeks, then 1 capsule at night for 2 weeks. Duration 13 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=22) Intervention 3: Placebo. Matching placebo. 1 capsule at night for 1 week, then 2 capsules daily for 10 weeks, then 1 capsule at night for 2 weeks. Duration 13 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Academic or government funding (This work was supported by The Rosetrees' Trust, grant number M11-F1, by the UK National Institute of Health Clinical Research Network and NIHR clinical academic fellowship to MR.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PREGABALIN versus SNRIS

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN pain score at 13 weeks; Group 1: mean -132.1 (SD 117.5); n=22, Group 2: mean -35.8 (SD 196.3); n=21; AUSCAN pain score 0-500 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -132.1 (-181.1 to -82.9). Reported duloxetine: -35.8 (-119.7 to 48.2). Calculated SD pregabalin: 117.5. Calculated SD duloxetine: 196.3. Baseline pregabalin: 317.0 (280.8 to 353.1). Baseline duloxetine: 296.0 (248.2 to 343.9).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN function score at 13 weeks; Group 1: mean -246.4 (SD 228.2); n=22, Group 2: mean -101.8 (SD 238.1); n=21; AUSCAN function score 0-900 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -246.4 (-

341.7 to -151.0). Reported duloxetine: -101.8 (-248.4 to -44.7). Calculated SD pregabalin: 228.2. Calculated SD duloxetine: 238.1. Baseline pregabalin: 576.2 (499.1 to 653.4). Baseline duloxetine: 577.2 (478.0 to 676.4).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.

Protocol outcome 3: Psychological distress at ≤3- or >3- months
- Actual outcome for Hand: Hospital Anxiety and Depression Score Anxiety subscale at 13 weeks; Group 1: mean -0.82 (SD 3.1); n=22, Group 2: mean -1.3 (SD 4.2); n=21; Hospital Anxiety and Depression Score Anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -0.82 (-2.1 to 0.5). Reported duloxetine: -1.3 (-3.1 to 0.5). Calculated SD pregabalin: 3.1. Calculated SD duloxetine: 4.2. Baseline pregabalin: 6.5 (4.6 to 8.4). Baseline duloxetine: 5.9 (4.3 to 7.6).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.
- Actual outcome for Hand: Hospital Anxiety and Depression Score Depression subscale at 13 weeks; Group 1: mean -1.1 (SD 2.5); n=22, Group 2: mean -0.3 (SD 3.6); n=21; Hospital Anxiety and Depression Score Depression subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -1.1 (-2.1 to -0.02). Reported duloxetine: -0.3 (-1.9 to 1.2). Calculated SD pregabalin: 2.5. Calculated SD duloxetine: 3.6. Baseline pregabalin: 5.1 (3.6 to 6.7). Baseline duloxetine: 4.4 (2.9 to 5.8).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hand: Digestive adverse events at 13 weeks; Group 1: 7/22, Group 2: 18/21
Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.

Protocol outcome 5: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Hand: Cardiovascular adverse events at 13 weeks; Group 1: 3/22, Group 2: 2/21

Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data – High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PREGABALIN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN pain score at 13 weeks; Group 1: mean -132.1 (SD 117.5); n=22, Group 2: mean -46.61 (SD 113.3); n=22; AUSCAN pain score 0-500 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -132.1 (-181.1 to -82.9). Reported placebo: -46.61 (-93.9 to 0.75). Calculated SD pregabalin: 117.5. Calculated SD placebo: 113.3. Baseline pregabalin: 317.0 (280.8 to 353.1). Baseline placebo: 320.3 (290.9 to 349.6).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN function score at 13 weeks; Group 1: mean -246.4 (SD 228.2); n=22, Group 2: mean -67.3 (SD 161.1); n=22; AUSCAN function score 0-900 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -246.4 (-341.7 to -151.0). Reported placebo: -67.3 (-156.4 to -21.8). Calculated SD pregabalin: 228.2. Calculated SD placebo: 161.1. Baseline pregabalin: 576.2 (499.1 to 653.4). Baseline placebo: 582.3 (509.1 to 655.5).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 3: Psychological distress at ≤3- or >3- months
- Actual outcome for Hand: Hospital Anxiety and Depression Score Anxiety subscale at 13 weeks; Group 1: mean -0.82 (SD 3.1); n=22, Group 2: mean 0.5 (SD 2.2); n=22; Hospital Anxiety and Depression score anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -0.82 (-2.1 to 0.5). Reported placebo: 0.5 (-0.4 to 1.4). Calculated SD pregabalin: 3.1. Calculated SD placebo: 2.2. Baseline pregabalin: 6.5 (4.6 to 8.4). Baseline placebo: 7.2 (5.4 to 9.0).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1

was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.
- Actual outcome for Hand: Hospital Anxiety and Depression Score Depression subscale at 13 weeks; Group 1: mean -1.1 (SD 2.5); n=22, Group 2: mean 0.05 (SD 3.2); n=22; Hospital Anxiety and Depression Score Depression subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -1.1 (-2.1 to -0.02). Reported placebo: 0.05 (-1.3 to 1.4). Calculated SD pregabalin: 2.5. Calculated SD placebo: 3.2. Baseline pregabalin: 5.1 (3.6 to 6.7). Baseline placebo: 4.9 (3.6 to 6.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hand: Digestive adverse events at 13 weeks; Group 1: 7/22, Group 2: 5/22
Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 5: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Hand: Cardiovascular adverse events at 13 weeks; Group 1: 3/22, Group 2: 1/22
Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Pregabalin: 1 withdrew due to a family bereavement, 1 was noncompliant with medication, 1 lost to follow up, 2 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN pain score at 13 weeks; Group 1: mean -35.8 (SD 196.3); n=21, Group 2: mean -46.61 (SD 113.3); n=22; AUSCAN pain score 0-500 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported duloxetine: -35.8 (-119.7 to 48.2). Reported placebo: -46.61 (-93.9 to 0.75). Calculated SD duloxetine: 196.3. Calculated SD placebo: 113.3. Baseline duloxetine: 296.0 (248.2 to 343.9). Baseline placebo: 320.3 (290.9 to 349.6).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4

withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Hand: AUSCAN function score at 13 weeks; Group 1: mean -101.8 (SD 238.1); n=21, Group 2: mean -67.3 (SD 161.1); n=22; AUSCAN function score 0-900 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported pregabalin: -246.4 (-341.7 to -151.0). Reported duloxetine: -101.8 (-248.4 to -44.7). Reported placebo: -67.3 (-156.4 to -21.8). Calculated SD pregabalin: 228.2. Calculated SD duloxetine: 238.1. Calculated SD placebo: 161.1. Baseline pregabalin: 576.2 (499.1 to 653.4). Baseline duloxetine: 577.2 (478.0 to 676.4). Baseline placebo: 582.3 (509.1 to 655.5).
Risk of bias: All domain – High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 3: Psychological distress at ≤3- or >3- months
- Actual outcome for Hand: Hospital Anxiety and Depression Score Anxiety subscale at 13 weeks; Group 1: mean -1.3 (SD 4.2); n=21, Group 2: mean 0.5 (SD 2.2); n=22; Hospital Anxiety and Depression Score Anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported duloxetine: -1.3 (-3.1 to 0.5). Reported placebo: 0.5 (-0.4 to 1.4). Calculated SD duloxetine: 4.2. Calculated SD placebo: 2.2. Baseline duloxetine: 5.9 (4.3 to 7.6). Baseline placebo: 7.2 (5.4 to 9.0).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.
- Actual outcome for Hand: Hospital Anxiety and Depression Score Depression subscale at 13 weeks; Group 1: mean -0.3 (SD 3.6); n=21, Group 2: mean 0.05 (SD 3.2); n=22; Hospital Anxiety and Depression Score Depression subscale 0-21 Top=High is poor outcome; Comments: Reports change score and 95% confidence interval. Reported duloxetine: -0.3 (-1.9 to 1.2). Reported placebo: 0.05 (-1.3 to 1.4). Calculated SD duloxetine: 3.6. Calculated SD placebo: 3.2. Baseline duloxetine: 4.4 (2.9 to 5.8). Baseline placebo: 4.9 (3.6 to 6.2).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Hand: Digestive adverse events at 13 weeks; Group 1: 18/21, Group 2: 5/22
Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.

Protocol outcome 5: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Hand: Cardiovascular adverse events at 13 weeks; Group 1: 2/21, Group 2: 1/22 Risk of bias: All domain – Very high, Selection - Low, Blinding - Low, Incomplete outcome data – High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, BMI, most common analgesics before inclusion and baseline values of outcomes; Group 1 Number missing: 5, Reason: Duloxetine: 1 developed bronchitis and withdrew, 4 withdrew due to adverse events.; Group 2 Number missing: 3, Reason: Placebo: 2 lost to follow up, 1 withdrew due to adverse events.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Strand 2017[184]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=305)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographic evidence of hip and/or knee osteoarthritis
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women at least 40 years old, weight at least 45kg, with radiographic evidence of osteoarthritis of the hip and/or knee who were receiving chronic treatment with NSAIDs and/or paracetamol. Eligible people had baseline WOMAC pain subscale scores of at least 40mm by a 100mm visual analogue scale with a documented osteoarthritis flare, defined as a at least 15mm increase in WOMAC pain subscale score from screening to baseline following NSAID or paracetamol discontinuation (at least 5 washout days).
Exclusion criteria	History of allergic reaction or clinically significant intolerance to paracetamol, aspirin, or any NSAIDs; requires continuous use of opioid or opioid combination products to control osteoarthritis pain of the knee or hip; clinically significant unstable cardiac, respiratory, neurological, immunological, haematological, or renal disease; significant difficulties swallowing capsules or unable to tolerate oral medication; previous participation in another clinical study of diclofenac capsules or received any investigational drug or device or investigational therapy within 30 days before screening
Recruitment/selection of patients	Requires people to meet a flare criteria after discontinuation of NSAIDs or paracetamol
Age, gender and ethnicity	Age - Mean (SD): 61.6 (8.90). Gender (M:F): 102:203. Ethnicity: White = 245, Black or African American = 57, Native Hawaiian or other Pacific Islander = 1, Native American or Alaskan Native = 1, Other = 2
Further population details	1. Age: Mixed (Based on range: 41-90). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).

Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated. NCT01461369
Indirectness of population	No indirectness
Interventions	(n=202) Intervention 1: NSAIDs - Diclofenac. SoluMatrix diclofenac capsules, 35mg three times daily or twice daily. Duration 12 weeks. Concurrent medication/care: Paracetamol was permitted as a rescue medication during the washout and treatment periods (up to 3000mg daily) but was discouraged within 24 hours and prohibited within 12 hours prior to scheduled study visits on day 0 and at weeks 2, 6 and 12 during the treatment period. Indirectness: No indirectness Comments: These two groups were combined due to class effect as agreed in the protocol (n=103) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: Paracetamol was permitted as a rescue medication during the washout and treatment periods (up to 3000mg daily) but was discouraged within 24 hours and prohibited within 12 hours prior to scheduled study visits on day 0 and at weeks 2, 6 and 12 during the treatment period. Indirectness: No indirectness
Funding	Study funded by industry (Funding was provided by Iroko)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 physical functioning subscale at 12 weeks; Group 1: mean 14.5 (SD 22.54); n=202, Group 2: mean 7.5 (SD 22.94); n=103; SF-36 physical function subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 14.7 (2.31). Reported diclofenac 35mg BD (n=104): 14.3 (2.18). Reported placebo (n=103): 7.5 (2.26). Baseline diclofenac 35mg TDS: 35.3 (20.35). Baseline diclofenac 35mg BD: 35.7 (19.52). Baseline placebo: 36.2 (21.41).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 role physical subscale at 12 weeks; Group 1: mean 17.2 (SD 24.57); n=202, Group 2: mean 11 (SD 24.97); n=103; SF-36 role physical subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 18.1 (2.52). Reported diclofenac 35mg BD (n=104): 16.3 (2.37). Reported placebo (n=103): 11.0 (2.46). Baseline diclofenac 35mg TDS: 41.2 (23.39). Baseline diclofenac 35mg BD: 42.8 (22.73). Baseline placebo: 44.0 (22.59).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 bodily pain subscale at 12 weeks; Group 1: mean 21.9 (SD 21.94); n=202, Group 2: mean 12.8 (SD 22.23); n=103; SF-36 bodily pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 21.8 (2.25). Reported diclofenac 35mg BD (n=104): 22.0 (2.12). Reported placebo (n=103): 12.8 (2.19). Baseline diclofenac 35mg TDS: 35.0 (14.27). Baseline diclofenac 35mg BD: 33.8 (14.31). Baseline placebo: 33.5 (13.84).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 general health subscale at 12 weeks; Group 1: mean 2 (SD 17.23); n=202, Group 2: mean -0.1 (SD 17.46); n=103; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 2.7 (1.76). Reported diclofenac 35mg BD (n=104): 1.4 (1.67). Reported placebo (n=103): -0.1 (1.72). Baseline diclofenac 35mg TDS: 67.8 (20.49). Baseline diclofenac 35mg BD: 71.3 (18.74). Baseline placebo: 64.1 (21.88).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 vitality subscale at 12 weeks; Group 1: mean 8.9 (SD 17.54); n=202, Group 2: mean 3 (SD 17.66); n=103; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 9.8 (1.80). Reported diclofenac 35mg BD (n=104): 8.0 (1.69). Reported placebo (n=103): 3.0 (1.74). Baseline diclofenac 35mg TDS: 49.2 (19.88). Baseline diclofenac 35mg BD: 46.7 (19.37). Baseline placebo: 46.8 (18.43).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 social functioning subscale at 12 weeks; Group 1: mean 11.6 (SD 22.76); n=202, Group 2: mean 7 (SD 22.94); n=103; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 13.4 (2.33). Reported diclofenac 35mg BD (n=104): 9.8 (2.19). Reported placebo (n=103): 7.0 (2.26). Baseline diclofenac 35mg TDS: 66.5 (23.42). Baseline diclofenac 35mg BD: 65.9 (25.24). Baseline placebo: 65.0 (23.37).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports

overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 role emotional subscale at 12 weeks; Group 1: mean 9.3 (SD 24.47); n=202, Group 2: mean 7.2 (SD 25.17); n=103; SF-36 role emotional subscale 0-100 Top=High is good outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 10.3 (2.51). Reported diclofenac 35mg BD (n=104): 8.4 (2.36). Reported placebo (n=103): 7.2 (2.48). Baseline diclofenac 35mg TDS: 67.8 (27.52). Baseline diclofenac 35mg BD: 68.4 (28.80). Baseline placebo: 62.2 (25.39).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up
- Actual outcome for Other: SF-36 mental health subscale at 12 weeks; Group 1: mean 4.3 (SD 16.37); n=202, Group 2: mean 1.9 (SD 16.64); n=103; SF-36 mental health subscale 0-100 Top=High is poor outcome; Comments: Reports least square means and standard error. Converted to standard deviation. Reported diclofenac 35mg TDS (n=98): 3.8 (1.68). Reported diclofenac 35mg BD (n=104): 4.7 (1.58). Reported placebo (n=103): 1.9 (1.64). Baseline diclofenac 35mg TDS: 75.7 (16.67). Baseline diclofenac 35mg BD: 75.0 (17.72). Baseline placebo: 72.5 (17.50).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -41.3 (SD 31.5); n=202, Group 2: mean -32.2 (SD 31.36); n=103; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports individual question answers as least square means and standard error. These were converted to standard deviations and then pooled to determine the total subscale value (on a 0-100 scale). Reported diclofenac TDS 1A (n=98): -44.1 (3.07). Reported diclofenac TDS 1B: -44.7 (3.17). Reported diclofenac TDS 1C: -41.0 (3.36). Reported diclofenac TDS 1D: -43.0 (3.26). Reported diclofenac TDS 1E: -44.0 (3.22). Reported diclofenac TDS 1F: -45.8 (3.15). Reported diclofenac BD 1A (n=104): -39.0 (2.91). Reported diclofenac BD 1B: -40.0 (3.00). Reported diclofenac BD 1C: -36.0 (3.19). Reported diclofenac BD 1D: -38.2 (3.08). Reported diclofenac BD 1E: -38.4 (3.04). Reported diclofenac BD 1F: -42.2 (2.98). Reported placebo 1A (n=103): -32.5 (2.94). Reported placebo 1B: -31.6 (3.03). Reported placebo 1C: -26.6 (3.26). Reported placebo 1D: -33.6 (3.13). Reported placebo 1E: -34.1 (3.08). Reported placebo 1F: -34.6 (3.02). Calculated SD diclofenac TDS 1A: 30.39. Calculated SD diclofenac TDS 1B: 31.38. Calculated SD diclofenac TDS 1C: 33.26. Calculated SD diclofenac TDS 1D: 32.27. Calculated SD diclofenac TDS 1E: 31.88. Calculated SD diclofenac TDS 1F: 31.18. Calculated SD diclofenac BD 1A: 29.68. Calculated SD diclofenac BD 1B: 30.59. Calculated SD diclofenac BD 1C: 32.53. Calculated SD diclofenac BD 1D: 31.41. Calculated SD diclofenac BD 1E: 31.41. Calculated SD diclofenac BD 1F: 30.39. Calculated SD placebo 1A: 29.84. Calculated SD placebo 1B: 30.75 . Calculated SD placebo 1C: 33.09. Calculated SD placebo 1D: 31.77. Calculated SD placebo 1E: 31.26. Calculated SD placebo 1F: 30.65. Pooled score diclofenac TDS: -43.8 (31.77). Pooled scored diclofenac BD: -39.0 (31.07). Pooled score placebo: -32.2 (31.36).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI, target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports

overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 12 weeks; Group 1: mean -33.1 (SD 46.97); n=202, Group 2: mean -24.4 (SD 46.47); n=103; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports individual question answers as least square means and standard error. These were converted to standard deviations and then pooled to determine the total subscale value (on a 0-100 scale). Reported diclofenac TDS 3A (n=98): -35.4 (3.15). Reported diclofenac TDS 3B: -32.2 (3.32). Reported diclofenac TDS 3C: -31.6 (3.35). Reported diclofenac TDS 3D: -37.5 (3.30). Reported diclofenac TDS 3E: -34.6 (3.19). Reported diclofenac TDS 3F: -35.3 (3.44). Reported diclofenac TDS 3G: -36.4 (3.31). Reported diclofenac TDS 3H: -38.4 (3.32). Reported diclofenac TDS 3I: -36.9 (3.27). Reported diclofenac TDS 3J: -35.3 (3.27). Reported diclofenac TDS 3K: -38.8 (3.21). Reported diclofenac TDS 3L: -36.4 (3.31). Reported diclofenac TDS 3M: -37.4 (3.14). Reported diclofenac TDS 3N: -34.1 (14.87). Reported diclofenac TDS 3O: -33.0 (3.10). Reported diclofenac TDS 3P: -37.3 (3.25). Reported diclofenac TDS 3Q: -33.5 (3.41). Reported diclofenac TDS 3R: -33.6 (3.19). Reported diclofenac BD 3A (n=104): -30.5 (2.98). Reported diclofenac BD 3B: -30.2 (3.17). Reported diclofenac BD 3C: -29.6 (3.19). Reported diclofenac BD 3D: -33.4 (3.13). Reported diclofenac BD 3E: -32.7 (3.03). Reported diclofenac BD 3F: -33.3 (3.26). Reported diclofenac BD 3G: -32.7 (3.13). Reported diclofenac BD 3H: -34.1 (3.15). Reported diclofenac BD 3I: -33.9 (3.09). Reported diclofenac BD 3J: -31.2 (3.10). Reported diclofenac BD 3K: -34.7 (3.03). Reported diclofenac BD 3L: -33.3 (3.13). Reported diclofenac BD 3M: -32.5 (2.98). Reported diclofenac BD 3N: -11.3 (14.63). Reported diclofenac BD 3O: -30.8 (2.93). Reported diclofenac BD 3P: -33.4 (3.08). Reported diclofenac BD 3Q: -30.8 (3.20). Reported diclofenac BD 3R: -29.8 (3.03). Reported placebo 3A (n=103): -24.2 (3.07). Reported placebo 3B: -20.6 (3.23). Reported placebo 3C: -21.2 (3.26). Reported placebo 3D: -24.2 (3.16). Reported placebo 3E: -24.2 (3.07). Reported placebo 3F: -25.1 (3.30). Reported placebo 3G: -25.0 (3.17). Reported placebo 3H: -24.6 (3.19). Reported placebo 3I: -24.5 (3.19). Reported placebo 3J: -23.7 (3.15). Reported placebo 3K: -28.1 (3.08). Reported placebo 3L: -25.3 (3.18). Reported placebo 3M: -28.8 (3.03). Reported placebo 3N: -23.6 (14.42). Reported placebo 3O: -24.7 (2.97). Reported placebo 3P: -26.1 (3.12). Reported placebo 3Q: -22.9 (3.30). Reported placebo 3R: -22.5 (3.06). Calculated SD diclofenac TDS 3A: 31.18. Calculated SD diclofenac TDS 3B: 32.87. Calculated SD diclofenac TDS 3C: 33.16. Calculated SD diclofenac TDS 3D: 32.67. Calculated SD diclofenac TDS 3E: 31.58. Calculated SD diclofenac TDS 3F: 34.05. Calculated SD diclofenac TDS 3G: 32.77. Calculated SD diclofenac TDS 3H: 32.87. Calculated SD diclofenac TDS 3I: 32.37. Calculated SD diclofenac TDS 3J: 32.37. Calculated SD diclofenac TDS 3K: 31.78. Calculated SD diclofenac TDS 3L: 32.77. Calculated SD diclofenac TDS 3M: 31.08. Calculated SD diclofenac TDS 3N: 147.20. Calculated SD diclofenac TDS 3O: 30.69. Calculated SD diclofenac TDS 3P: 32.17. Calculated SD diclofenac TDS 3Q: 33.76. Calculated SD diclofenac TDS 3R: 31.58. Calculated SD diclofenac BD 3A: 30.39. Calculated SD diclofenac BD 3B: 32.33. Calculated SD diclofenac BD 3C: 32.53. Calculated SD diclofenac BD 3D: 31.92. Calculated SD diclofenac BD 3E: 30.90. Calculated SD diclofenac BD 3F: 33.25. Calculated SD diclofenac BD 3G: 31.91. Calculated SD diclofenac BD 3H: 32.12. Calculated SD diclofenac BD 3I: 31.51. Calculated SD diclofenac BD 3J: 31.61. Calculated SD diclofenac BD 3K: 30.90. Calculated SD diclofenac BD 3L: 31.92. Calculated SD diclofenac BD 3M: 30.39. Calculated SD diclofenac BD 3N: 149.20. Calculated SD diclofenac BD 3O: 29.88. Calculated SD diclofenac BD 3P: 31.41. Calculated SD diclofenac BD 3Q: 32.63. Calculated SD diclofenac BD 3R: 30.90. Calculated SD placebo 3A: 31.16. Calculated SD placebo 3B: 32.78. Calculated SD placebo 3C: 33.09. Calculated SD placebo 3D: 32.07. Calculated SD placebo 3E: 31.16. Calculated SD placebo 3F: 33.49. Calculated SD placebo 3G: 32.17. Calculated SD placebo 3H: 32.38. Calculated SD placebo 3I: 32.38. Calculated SD placebo 3J: 31.97. Calculated SD placebo 3K: 31.26. Calculated SD placebo 3L: 32.27. Calculated SD placebo 3M: 30.75. Calculated SD placebo 3N: 146.35. Calculated SD placebo 3O: 30.14 . Calculated SD placebo 3P: 31.67. Calculated SD placebo 3Q: 33.49. Calculated SD placebo 3R: 31.06. Pooled score diclofenac TDS: -35.4 (46.90). Pooled scored diclofenac BD: -31.0 (46.94). Pooled score placebo: -24.4 (46.47).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, weight, height, BmI,

target osteoarthritis joint and baseline values of outcomes (apart from WOMAC physical function subscale); Group 1 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up; Group 2 Number missing: -, Reason: Reports overall rate only. 25 withdrew due to adverse events, 8 lack of efficacy, 6 withdrew consent, 5 protocol violations, 3 lost to follow up	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3-months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3-months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3-months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3-months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Tannenbaum 2004[185]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1702)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee according to the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women aged ≥18 years with a confirmed diagnosis of primary osteoarthritis of the knee, according to the American College of Rheumatology criteria.
Exclusion criteria	People at risk of pregnancy or those who had secondary osteoarthritis, other connective tissue diseases or significant medical problems
Recruitment/selection of patients	People meeting the initial inclusion criteria underwent a 3-7 day washout period, during which NSAID treatment was not permitted. At the end of the washout period, people with pain intensity in the affected knee measuring greater than or equal to 40mm on a 100mm visual analogue scale (most pain) in the past 24 hours were deemed eligible for entry into the treatment phase of the study. To best reflect the "real life" clinical situation, no increase/worsening in osteoarthritis symptoms (flare) was required for study entry
Age, gender and ethnicity	Age - Mean (SD): 64.2 (10.4). Gender (M:F): 536:1166. Ethnicity: White = 1684
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not reported Duration of symptoms [median]: 4.75 years
Indirectness of population	No indirectness
Interventions	(n=481) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg once a day. Duration 13 weeks. Concurrent medication/care: People were permitted to take paracetamol (≤2 grams/day). However, they were asked to refrain from using the rescue drug from midnight before each clinic visit. NSAIDs were not permitted during the course of the study, with the exception of low dose aspirin (≤325mg/day) for a cardiovascular indication. Indirectness: No indirectness

	(n=978) Intervention 2: NSAIDs - Other. Lumiracoxib 200mg once a day and lumiracoxib 400mg once a day. Duration 13 weeks. Concurrent medication/care: People were permitted to take paracetamol (≤2 grams/day). However, they were asked to refrain from using the rescue drug from midnight before each clinic visit. NSAIDs were not permitted during the course of the study, with the exception of low dose aspirin (≤325mg/day) for a cardiovascular indication. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK, therefore it was not included in the analysis. It was reported here for completeness. (n=243) Intervention 3: Placebo. Placebo once daily. Duration 13 weeks. Concurrent medication/care: People were permitted to take paracetamol (≤2 grams/day). However, they were asked to refrain from using the rescue drug from midnight before each clinic visit. NSAIDs were not permitted during the course of the study, with the exception of low dose aspirin (≤325mg/day) for a cardiovascular indication. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a grant from Novartis Pharma AG, Basel, Switzerland)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 13 weeks; Group 1: mean -3.1 (SD 3.8); n=481, Group 2: mean -2.4 (SD 3.8); n=243; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 10.1 (3.3). Baseline placebo: 10.3 (3.0).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, body mass index, OA disease duration and baseline values in outcomes; Group 1 Number missing: 80, Reason: 80 withdrew - 15 unsatisfactory therapeutic effect, 1 lost to follow up, 47 adverse events, 17 other; Group 2 Number missing: 43, Reason: 43 withdrew - 15 unsatisfactory therapeutic effect, 3 lost to follow up, 21 adverse events, 4 other

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 13 weeks; Group 1: mean -9.2 (SD 11.6); n=481, Group 2: mean -6.2 (SD 11.8); n=243; WOMAC DPDA subscale 0-68 Top=High is poor outcome; Comments: Baseline celecoxib: 34.4 (11.7). Baseline placebo: 34.6 (10.4).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, body mass index, OA disease duration and baseline values in outcomes; Group 1 Number missing: 80, Reason: 80 withdrew - 15 unsatisfactory therapeutic effect, 1 lost to follow up, 47

adverse events, 17 other; Group 2 Number missing: 43, Reason: 43 withdrew - 15 unsatisfactory therapeutic effect, 3 lost to follow up, 21 adverse events, 4 other Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal events (including nasopharyngitis, upper abdominal pain, dyspepsia and diarrhoea) at 13 weeks; Group 1: 72/481, Group 2: 25/243; Comments: Celecoxib: GI events (total): 72. Nasopharyngitis = 23, upper abdominal pain = 25, dyspepsia = 17, diarrhoea = 11. Placebo: GI events (total): 25. Nasopharyngitis = 12, upper abdominal pain = 6, dyspepsia = 9, diarrhoea = 2 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, body mass index, OA disease duration and baseline values in outcomes; Group 1 Number missing: 80, Reason: 80 withdrew - 15 unsatisfactory therapeutic effect, 1 lost to follow up, 47 adverse events, 17 other; Group 2 Number missing: 43, Reason: 43 withdrew - 15 unsatisfactory therapeutic effect, 3 lost to follow up, 21 adverse events, 4 other Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 13 weeks; Group 1: 27/481, Group 2: 9/243 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, ethnicity, body mass index, OA disease duration and baseline values in outcomes; Group 1 Number missing: 80, Reason: 80 withdrew - 15 unsatisfactory therapeutic effect, 1 lost to follow up, 47 adverse events, 17 other; Group 2 Number missing: 43, Reason: 43 withdrew - 15 unsatisfactory therapeutic effect, 3 lost to follow up, 21 adverse events, 4 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Temple 2006[186]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=581)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the hip or knee for at least 6 months with a history of pain of mild or moderate intensity on a 5 point scale at screening. Had to fulfill at least 2 of the following 5 criteria based, in part, on the clinical criteria established by the American Rheumatism Association: morning stiffness of <30 minutes' duration, crepitus, bony tenderness, bony enlargement, and no palpable warmth. Eligible people also had recent (obtain at least 6 months before screening) radiographic evidence of grade 2 or 3 osteoarthritis of the knee or hip (osteophytes at the joint margin), and were either American College of Rheumatology functional class I-III.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Symptomatic osteoarthritis of the hip or knee for at least 6 months with a history of pain of mild or moderate intensity on a 5 point scale at screening. They had to require treatment with either an analgesics or anti-inflammatory agent for at least 3 days per week for at least 3 months before enrollment. People with osteoarthritis of the knee had to fulfill at least 2 of the following 5 criteria based, in part, on the clinical criteria established by the American Rheumatism Association: morning stiffness of <30 minutes' duration, crepitus, bony tenderness, bony enlargement, and no palpable warmth. Eligible people also had recent (obtain at least 6 months before screening) radiographic evidence of grade 2 or 3 osteoarthritis of the knee or hip (osteophytes at the joint margin), and were either American College of Rheumatology functional class I-II. Laboratory values had to be consistent with a diagnosis of osteoarthritis, including an erythrocyte sedimentation rate of <40mm/h and a rheumatoid factor of <40 IU/mL.
Exclusion criteria	A history of recent (previous 12 months) surgery (including arthroscopy) or major trauma to the study joint; active inflammation; more severe radiographic criteria; other clinical or laboratory evidence of more serious joint disease; people with secondary osteoarthritis of the study joint or pseudogout of any joint; people requiring other long-term treatment with drugs that might interfere with assessments (e.g. aspirin doses >325mg/d); anticoagulants; corticosteroids within 2 months; hyaluronan within 3

	months; non-stable doses of anticonvulsants, antidepressants, tranquilizers, hypolipidaemic agents, glucosamine and chondroitin sulfate; pregnancy; lactation; previous gastric surgery; allergy to aspirin, NSAIDs or study medications; active illnesses that could interfere with the conduct of the study (e.g. peptic ulcer disease, inflammatory bowel disease, clinically important renal or hepatic disease, based on the investigator's judgement of the patient's clinical history or baseline laboratory assessments); drug dependency; abuse of drugs or alcohol (more than or equal to 3 drinks/day).
Recruitment/selection of patients	Eligible people were required to discontinue their current antiarthritis medication(s) for a washout period ranging from 3 to 11 days, depending on the medication half-life. At the baseline visit, people who reported mild, moderate or moderately severe pain intensity over the previous 24 hours and had an at least 20% increase in WOMAC pain subscale score (relative to screening) were enrolled in the study
Age, gender and ethnicity	Age - Mean (SD): 59.3 (8.6). Gender (M:F): 176:395. Ethnicity: White = 480, Black = 51, Hispanic = 38, Other = 2
Further population details	1. Age: <75 years (Range went up to 75 years). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Kellgren Lawrence grade 2-3. American College of Rheumatology functional class I-II. Duration of symptoms: Not stated explicitly. At least 3 months.
Indirectness of population	No indirectness
Interventions	(n=290) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 4 grams/day (1 gram every 4-6 hours). Duration 12 months (53 people only received treatment for 6 months after recruitment at a later date). Concurrent medication/care: No additional information. Indirectness: No indirectness (n=291) Intervention 2: NSAIDs - Naproxen. Naproxen 750mg/day (375mg twice a day).. Duration 12 months (52 people only received treatment for 6 months after recruitment at a later date). Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by McNeil Consumer & Specialty Pharmaceuticals (protocol 98-055). Dr Benson has served as a consultant to McNeil Consumer & Specialty Pharmaceuticals.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARACETEMOL versus NAPROXEN

Protocol outcome 1: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Other: Peripheral oedema at 12 months; Group 1: 3/287, Group 2: 11/284
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High, Other 2 - Low, Comments - Midway through the study they added additional participants; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, height and weight; Group 1 Number missing: 141, Reason: 237 people in the original study. Out of those 182 withdrew (58 adverse events, 62 for lack of efficacy, 11 lost to follow up, 8 protocol violations, 43 for other). Then at 6 months 53 added in, of these 17 withdrew (2 lack of efficacy, 2 lost to follow up, 4 protocol violations, 9 other).; Group 2 Number missing: 124, Reason: 239 people in the original study. Out of those 159 withdrew (55 adverse events, 52 lack of efficacy, 14 lost to follow up, 6 protocol violations, 32 others). 52 in the second group, of those 20 withdrew (3 lack of efficacy, 3 lost to follow up, 5 protocol violations, 9 others).

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Thorne 2008[187]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (4 weeks for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with osteoarthritis defined by the presence of hip and/or knee symptoms (pain, stiffness, disability) and signs (bony crepitus) as well as radiographic evidence of osteoarthritis in the medial and/or lateral tibiofemoral compartment (with or without patellofemoral osteoarthritis) or in the hip. Radiographic evidence was defined by the presence of at least one of the following: osteophytes, joint space narrowing, periarticular sclerosis or subchondral cysts, with a minimum grade 2 severity. People with more advanced grades were eligible if they were not awaiting surgery.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and nonpregnant, non-nursing women over the age of 18 years, diagnosed with osteoarthritis and requiring the use of paracetamol, anti-inflammatory agents or combination opioid and nonopioid analgesics for at least 3 months. People using only paracetamol at the time of enrollment were required to have pain of at least moderate intensity (a 2 or greater on a 0-4 ordinal pain scale) after a 2-7 day washout period at visit 2.
Exclusion criteria	People with intolerance to any opioid, tramadol or paracetamol; people who required more than eight tablets per day of paracetamol plus codeine, or its analgesic equivalent; people with a history of drug or alcohol abuse; people with the following medical conditions: any other form of joint disease or previous replacement of the study joint, renal or hepatic impairment (alanine aminotransferase or aspartate transferase more than 2 times the upper limit of the normal range), shortened gastrointestinal transit time, peptic ulcer disease, inflammatory disease of the gastrointestinal tract, cardiac or respiratory conditions that put the person at risk for respiratory depression, a history of seizures or a recognised risk for seizure, and any other condition that would adversely affect the person's safety or obscure the assessment of efficacy; people receiving monoamine oxidase inhibitors, carbamazepine, quinidine, selective serotonin reuptake inhibitors or tricyclic

	antidepressants, cyclobenzaprine, promethazine, neuroleptics, warfarin or digoxin; people who received an investigational drug within the last month.
Recruitment/selection of patients	If people were taking paracetamol before enrollment, they were required to have pain of at least moderate intensity (a 2 or greater on a 0-4 ordinal pain scale) after a 2 to 7 days washout period
Age, gender and ethnicity	Age - Mean (SD): 61.0 (10.3). Gender (M:F): 45:55. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: OA grade 2-3 Duration of symptoms (mean [SD]): 8.3 (6.8) years
Indirectness of population	No indirectness
Interventions	(n=100) Intervention 1: Strong opioids (oral) - Tramadol. Tramadol 150mg once daily uptitrated to the maxximal tolerable dose. Duration 4 weeks. Concurrent medication/care: Breakthrough pain was managed throughout the study with 325mg to 650mg plain paracetamol every 4 to 6 hours as required. Indirectness: No indirectness (n=100) Intervention 2: Placebo. Placebo tablet once per day. Duration 4 weeks. Concurrent medication/care: Breakthrough pain was managed throughout the study with 325mg to 650mg plain paracetamol every 4 to 6 hours as required. Indirectness: No indirectness
Funding	Study funded by industry (Supported by a research grant from Purdue Pharma, Canada)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: TRAMADOL versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 pain index at 4 weeks; Group 1: mean 38.81 (SD 10.76); n=94, Group 2: mean 35.61 (SD 9.01); n=88; SF-36 bodily pain subscale 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 16, Reason: 12 adverse events, 1 lack of efficacy, 1 consent withdrawn, 1 lost to follow-up, 1 other; Group 2 Number missing: 9, Reason: 3 adverse events, 3 lack of efficacy, 2 consent withdrawn, 1 protocol violation
- Actual outcome for Other: SF-36 general health perception at 4 weeks; Group 1: mean 46.54 (SD 11.2); n=94, Group 2: mean 44.39 (SD 11.63); n=88; SF-36 general health subscale 0-100 Top=High is good outcome

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 16, Reason: 12 adverse events, 1 lack of efficacy, 1 consent withdrawn, 1 lost to follow-up, 1 other; Group 2 Number missing: 9, Reason: 3 adverse events, 3 lack of efficacy, 2 consent withdrawn, 1 protocol violation - Actual outcome for Other: SF-36 vitality at 4 weeks; Group 1: mean 43.14 (SD 13.2); n=94, Group 2: mean 40.21 (SD 13.7); n=88; SF-36 vitality subscale 0-100 Top=High is good outcome Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 16, Reason: 12 adverse events, 1 lack of efficacy, 1 consent withdrawn, 1 lost to follow-up, 1 other; Group 2 Number missing: 9, Reason: 3 adverse events, 3 lack of efficacy, 2 consent withdrawn, 1 protocol violation Protocol outcome 2: Pain reduction at ≤3- or >3- months - Actual outcome for Other: WOMAC pain subscale at 4 weeks; Group 1: mean 189 (SD 105); n=94, Group 2: mean 230 (SD 115.4); n=88; WOMAC pain subscale 0-500 Top=High is poor outcome Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 16, Reason: 12 adverse events, 1 lack of efficacy, 1 consent withdrawn, 1 lost to follow-up, 1 other; Group 2 Number missing: 9, Reason: 3 adverse events, 3 lack of efficacy, 2 consent withdrawn, 1 protocol violation Protocol outcome 3: Physical function at ≤3- or >3- months - Actual outcome for Other: WOMAC physical function subscale at 4 weeks; Group 1: mean 632.4 (SD 361.3); n=94, Group 2: mean 727.4 (SD 383.4); n=88; WOMAC physical function subscale 0-1700 Top=High is poor outcome Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values only; Group 1 Number missing: 16, Reason: 12 adverse events, 1 lack of efficacy, 1 consent withdrawn, 1 lost to follow-up, 1 other; Group 2 Number missing: 9, Reason: 3 adverse events, 3 lack of efficacy, 2 consent withdrawn, 1 protocol violation	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Tiso 2010[188]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=20)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Partially adequate method of assessment/diagnosis: People of at least 50 years of age with at least 3 months of knee pain
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of at least 50 years old with at least 3 months of knee pain, willing and able to cooperate in the assigned treatment and willing and able to complete follow-up questionnaires
Exclusion criteria	Any contraindication to ibuprofen usage (known allergy or hypersensitivity, history of gastrointestinal ulcer or bleeding, renal disease, liver dysfunction, consuming >3 alcoholic beverages daily, or anticoagulant treatment); pregnancy; scheduled knee procedure within the 2 week follow up period
Recruitment/selection of patients	Any people taking ibuprofen or other NSAID at the time of enrollment underwent a 2-day washout period before beginning the study treatment. Other medications taken prior to the study were maintained, including those for pain management other than NSAIDs
Age, gender and ethnicity	Age - Mean (SD): 57.9 (9.2). Gender (M:F): 2:17. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Chronic pain grade I-IV - Median grade III Duration of symptoms: 3->12 months - median >12 months
Indirectness of population	Serious indirectness: Diagnostic criteria was too broad and could have included people with other rheumatological conditions
Interventions	(n=10) Intervention 1: NSAID gels (topical - local) - Ibuprofen gel. 4% ibuprofen gel - apply 2mL to the targeted area for an approximately 3.5cm diameter circle, 4 times daily (320mg total). Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness

	(n=10) Intervention 2: NSAIDs - Ibuprofen. Ibuprofen 800mg three times daily (2400mg total). Duration 2 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (Funding support was provided by Helm Pharmaceuticals)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN GEL versus IBUPROFEN

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: SF-12 physical component summary at 2 weeks; Group 1: mean 2.6 (SD 5.4); n=9, Group 2: mean 1.1 (SD 6.2); n=10; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reports mean differences and 95% confidence intervals. Converted to standard deviations. Reported ibuprofen gel: 2.6 (-0.9 to 6.2). Reported oral ibuprofen: 1.1 (-2.8 to 4.9). Baseline ibuprofen gel: 30.3 (8.1). Baseline oral ibuprofen: 30.5 (4.4).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, pain duration, chronic pain grade and baseline values of outcomes. WOMAC physical function and SF-12 mental component summary were significantly different at baseline.; Group 1 Number missing: 1, Reason: 1 lost to follow up; Group 2 Number missing: 0
- Actual outcome for Knee: SF-12 mental component summary at 2 weeks; Group 1: mean 5.1 (SD 31.1); n=9, Group 2: mean 3.9 (SD 9.4); n=10; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reports mean differences and 95% confidence intervals. Converted to standard deviations. Reported ibuprofen gel: 5.1 (-7.7 to 17.8). Reported oral ibuprofen: 3.9 (-2.0 to 9.7). Baseline ibuprofen gel: 33.7 (11.4). Baseline oral ibuprofen: 44.2 (13.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, pain duration, chronic pain grade and baseline values of outcomes. WOMAC physical function and SF-12 mental component summary were significantly different at baseline.; Group 1 Number missing: 1, Reason: 1 lost to follow up; Group 2 Number missing: 0

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 2 weeks; Group 1: mean -82.6 (SD 115.9); n=9, Group 2: mean -84.3 (SD 102.6); n=10; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Reports mean differences and 95% confidence intervals. Converted to standard deviations. Reported ibuprofen gel: -82.6 (-158.3 to -6.8). Reported oral ibuprofen: -84.3 (-177.9 to 9.3). Baseline ibuprofen gel: 305.6 (96.6). Baseline oral ibuprofen: 291.7 (83.3).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, pain duration, chronic pain grade and baseline values of outcomes. WOMAC physical function and SF-12 mental component summary were significantly different at baseline.; Group 1 Number missing: 1, Reason: 1 lost to follow up; Group 2 Number missing: 0

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 2 weeks; Group 1: mean -312.1 (SD 410.8); n=9, Group 2: mean -323.2 (SD 506.6); n=10;

WOMAC physical function subscale 0-1400 Top=High is poor outcome; Comments: Reports mean differences and 95% confidence intervals. Converted to standard deviations. Reported ibuprofen gel: -312.1 (-580.5 to -43.7). Reported oral ibuprofen: -323.2 (-637.2 to -9.2). Baseline ibuprofen gel: 1219.6 (201.4). Baseline oral ibuprofen: 1108.3 (302.7) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, pain duration, chronic pain grade and baseline values of outcomes. WOMAC physical function and SF-12 mental component summary were significantly different at baseline.; Group 1 Number missing: 1, Reason: 1 lost to follow up; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Trnavsky 2004[189]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=50)
Countries and setting	Conducted in Czech Republic; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis, unilateral or bilateral (ICD-10: M17.0/M17.1) as classified by the American College of Rheumatology criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People of any sex aged ≥40 and ≤75 years with primary knee osteoarthritis, unilateral or bilateral, chronic and decompensated, but not activated or without effusion or swelling, diagnosed according to the classification criteria of the American College of Rheumatology; knee pain on most days (>15) in the preceding month; radiographic osteophytes and grade II-III on the Kellgren-Lawrence OA severity score; score for pain on motion of ≥40mm on a 100mm visual analogue scale; total score ≥5 and ≤13 on the Lequesne algofunctional index; providing informed consent.
Exclusion criteria	Secondary osteoarthritis; obesity (body mass index ≥30kg/m²); chronic painful disease of the hip or the ankle joint; allergic diathesis, bronchial asthma, or known hypersensitivity to NSAID; eczematous skin eruption; any physiotherapy
Recruitment/selection of patients	A washout period of 1 to 60 days was mandatory to remove any type of pretreatment
Age, gender and ethnicity	Age - Mean (SD): 67.0 (7.1). Gender (M:F): 11:39. Ethnicity: Not stated
Further population details	1. Age: Mixed (Based on SD). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=25) Intervention 1: NSAID gels (topical - local) - Ibuprofen gel. Ibuprofen cream (5%) - 10cm cream three times daily (equivalent to 3x200mg ibuprofen).. Duration 8 days. Concurrent medication/care: Any concomitant treatment with other topical, intraarticular or systemic steroidal or nonsteroidal antirheumatics, analgesics, or

	disease modifying antirheumatic drugs was not allowed. However, it was assured that people received any medically necessary treatment (e.g. antihypertensives). During the washout period, peripherally-acting oral analgesics such as paracetamol were allowed as rescue medication up to 2 days before the start of the study treatment. Indirectness: No indirectness (n=25) Intervention 2: Placebo. Placebo cream, one 10cm strip 3 times daily. Duration 8 days. Concurrent medication/care: Any concomitant treatment with other topical, intraarticular or systemic steroidal or nonsteroidal antirheumatics, analgesics, or disease modifying antirheumatic drugs was not allowed. However, it was assured that people received any medically necessary treatment (e.g. antihypertensives). During the washout period, peripherally-acting oral analgesics such as paracetamol were allowed as rescue medication up to 2 days before the start of the study treatment. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Dolorgiet Pharmaceuticals, St. Augustin/Bonn, Germany)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN GEL versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain on motion (visual analogue scale) at 1 week; Group 1: mean 31.72 (SD 15.01); n=25, Group 2: mean 52.56 (SD 13.02); n=25; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline ibuprofen: 63.08 (7.27). Baseline placebo: 59.48 (7.98).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI and pain on motion; Group 1 Number missing: 12, Reason: 25 were enrolled. 12 made the 1st interim analysis (4 days) while 13 made the 2nd interim analysis (8 days). 25 in total were included in the analysis with all providing at least 1 post-baseline value.; Group 2 Number missing: 13, Reason: 25 were enrolled. 13 made the 1st interim analysis (4 days) while 12 made the 2nd interim analysis (8 days). 25 in total were included in the analysis with all providing at least 1 post-baseline value.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Adverse events at 1 week; Group 1: 0/25, Group 2: 0/25; Comments: "None of the 50 randomised patients experienced any adverse events during the study"
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI and pain on motion; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months

- Actual outcome for Knee: Adverse events at 1 week; Group 1: 0/25, Group 2: 0/25; Comments: "None of the 50 randomised patients experienced any adverse events during the study" Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI and pain on motion; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Adverse events at 1 week; Group 1: 0/25, Group 2: 0/25; Comments: "None of the 50 randomised patients experienced any adverse events during the study" Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI and pain on motion; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Adverse events at 1 week; Group 1: 0/25, Group 2: 0/25; Comments: "None of the 50 randomised patients experienced any adverse events during the study" Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, height, weight, BMI and pain on motion; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Trudeau 2015[190]
Study type	RCT (Patient randomised; Crossover: 1 week)
Number of studies (number of participants)	1 (n=63)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 week (2 phases)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary diagnosis of American College of Rheumatology Functional Class I to III osteoarthritis of the knee
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females over 21 years of age with a primary diagnosis of osteoarthritis of the knee if they had an in-clinic pain intensity (24-hour recall) ≤6 on a 0-10 numerical rating scale at screening on nonsteroidal anti-inflammatory drugs (NSAIDs), were able to withdraw osteoarthritis medications, and had a Hospital Anxiety and Depression Scale score ≤10 on either subscale (anxiety or depression).
Exclusion criteria	History of major depressive disorder not controlled with medication; people not meeting the flare criteria
Recruitment/selection of patients	Enrolled people entered a washout period in which they discontinued all analgesics for 1 week; paracetamol was allowed as rescue medication as needed (500mg maximum four times per day) during this period. In order to be randomised, people had to have an in-clinic pain intensity score (24-hour recall) of more than or equal to 4/10 on NRS and greater than or equal to 1 point higher than the screening score, as well as a Patient Global Assessment of Arthritis score that was worse than screening by greater than or equal to 1 category. People not meeting these criteria were discontinued from the study. People had to meet these criteria before randomisation and after the crossover and subsequent washout phase.
Age, gender and ethnicity	Age - Other: Older than 21 years, otherwise not reported. Gender (M:F): Not reported. Ethnicity: Not stated
Further population details	1. Age: Not stated / Unclear 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional Class I-III Duration of symptoms: Not stated

Indirectness of population	No indirectness
Interventions	(n=63) Intervention 1: NSAIDs - Celecoxib. Celecoxib 100mg twice a day for 1 week. Duration 1 week. Concurrent medication/care: Paracetamol was allowed as rescue medication as needed (500mg maximum four times a day). Indirectness: No indirectness (n=63) Intervention 2: Placebo. Placebo twice daily for 1 week. Duration 1 week. Concurrent medication/care: Paracetamol was allowed as rescue medication as needed (500mg maximum four times a day). Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by Pfizer, Inc.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO Protocol outcome 1: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function at 1 week; MD; 1.5 (Standard error: 6.23) WOMAC physical function 0-68 Top=High is poor outcome, Comments: Baseline celecoxib-placebo group: 14.5 (4.27). Baseline placebo-celecoxib group 14.2 (2.77).; Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports in-clinic pain, patient global assessment of arthritis, medial outcomes study sleep score, and WOMAC score and subscales. Does not report basic patient characteristics (ex. age, severity, duration of osteoarthritis symptoms, gender).; Group 1 Number missing: 10, Reason: 62 started the study. 52 finished the study. Reasons for exclusion were early discontinuation or not reaching the flare criteria for the second phase.; Group 2 Number missing: 7, Reason: 62 started the study. 55 finished the study. Reasons for exclusion were early discontinuation or not reaching the flare criteria for the second phase.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Truitt 2001[191]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=341)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fulfilling clinical and radiographic criteria for the diagnosis of tibiofemoral joint osteoarthritis of either knee or osteoarthritis of the hip
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 80 years or older if they met both clinical and radiographic criteria for the diagnosis of ostearthritis of either the knee or hip. If both knees and hips were affected, the most painful joint was designated the study joint and evaluated for patient inclusion and response to therapy. They required: pain in the study joint to be present for at least 6 months prior to screening, and clinical symptoms in the study joint had to be confirmed by the presence of radiographic findings (joint space narrowing in a hip, and joint space narrowing and osteophytes in a knee). All people were American College of Rheumatology Functional Classes I-III. People were required to have a history of positive therapeutic benefit from NSAIDs (including salicylates) or paracetamol, and to have taken this therapy on 20 or more of the previous 30 days. People were required to score 24 or more on the 30-question mini-mental status examination at screening. All people were required to swallow a test dose of placebo without substantial difficulty.
Exclusion criteria	Prior history of inflammatory arthritis (including rheumatoid arthritis), acute ligamentous or meniscal injury to the study joint within the past 18 months, or arthroscopy within 4 months of screening. People who received intraarticular or systemic corticosteroids within 3 months of entry were excluded. People with other medical conditions or laboratory abnormalities which either contraindicated the use of NSAIDs or were potential confounders of the safety evaluation were also excluded including: people with angina or congestive heart failure with symptoms at rest; people with serum creatinine greater than 2.0mg/dL or estimated creatinine clearances ≤30mL/min; uncontrolled hypertension; active gastrointestinal bleeding within the past three months; a history of leukaemia, lymphoma, or myeloproliferative disease; or

	hypersensitivity to aspirin or NSAIDs. People were excluded if any one of three pre-allocation stool-guaiac tests were positive.
Recruitment/selection of patients	The study was conducted in 48 community and university outpatient centers in the United States. All people were ambulatory with none residing in nursing homes.
Age, gender and ethnicity	Age - Mean (range): 83.0 (80-95). Gender (M:F): 124:217. Ethnicity: Not stated
Further population details	1. Age: ≥75 years (Only included people over the age of 80 years). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear (Potentially high comorbidity score. Some participants had hypertension, hypothyroidism, history of gastroduodenal ulcer or GI bleed, angina, myocardial infarction, coronary artery disease and drug allergy, and osteoporosis.). 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American College of Rheumatology Functional Classes I-III. Duration of symptoms: 15.0 (12.3) years.
Indirectness of population	No indirectness
Interventions	(n=115) Intervention 1: NSAIDs - Other. Oral nabumetone 1500mg once daily. Duration 6 weeks. Concurrent medication/care: Ongoing use of low dose aspirin (up to 325mg daily) was permitted. People were provided with paracetamol tablets (325mg) as rescue medication for breakthrough osteoarthritis pain. People were otherwise allowed to continue their usual therapy (including antacids, ACE inhibitors, beta blockers, diuretics, calcium antagonists, thyroid therapy and vitamin E). Indirectness: No indirectness (n=52) Intervention 2: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: Ongoing use of low dose aspirin (up to 325mg daily) was permitted. People were provided with paracetamol tablets (325mg) as rescue medication for breakthrough osteoarthritis pain. People were otherwise allowed to continue their usual therapy (including antacids, ACE inhibitors, beta blockers, diuretics, calcium antagonists, thyroid therapy and vitamin E). Indirectness: No indirectness (n=174) Intervention 3: NSAIDs - Other. Oral rofecoxib 12.5mg or 25mg once daily. Duration 6 weeks. Concurrent medication/care: Ongoing use of low dose aspirin (up to 325mg daily) was permitted. People were provided with paracetamol tablets (325mg) as rescue medication for breakthrough osteoarthritis pain. People were otherwise allowed to continue their usual therapy (including antacids, ACE inhibitors, beta blockers, diuretics, calcium antagonists, thyroid therapy and vitamin E). Indirectness: No indirectness

	Comments: Rofecoxib is not licensed for use in the UK so will not be included in the analysis. This is reported for completeness.
Funding	Study funded by industry (This study was funded by Merck Research Laboratories, Merck & Co., Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NABUMETONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale (100mm VAS) at 6 weeks; Group 1: mean -13.69 (SD 31); n=115, Group 2: mean -5.02 (SD 26.7); n=52; WOMAC pain subscale (100mm VAS) 0-100 Top=High is poor outcome; Comments: Reports least square mean change from baseline and 95% confidence intervals. Calculated SD from this. Reported nabumetone: -13.69 (-19.36, -8.03). Reported placebo: -5.02 (-12.27, 2.22). Baseline values not reported.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, duration of osteoarthritis, primary study joint, pre-study therapy, and mini-mental state examination score. Does not report baseline values for outcomes.; Group 1 Number missing: 15, Reason: Adverse events: 8. Lack of efficacy: 2. Other: 5 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).; Group 2 Number missing: 9, Reason: Adverse events: 1. Lack of efficacy: 6. Other: 2 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale (100mm VAS) at 6 weeks; Group 1: mean -14.42 (SD 28); n=115, Group 2: mean -5.51 (SD 24.1); n=52; WOMAC physical function subscale (100mm VAS) 0-100 Top=High is poor outcome; Comments: Reports least square mean change from baseline and 95% confidence intervals. Calculated SD from this. Reported nabumetone: -14.42 (-19.53, -9.31). Reported placebo: -5.51 (-12.05, 1.04). Baseline values not reported.
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, duration of osteoarthritis, primary study joint, pre-study therapy, and mini-mental state examination score. Does not report baseline values for outcomes.; Group 1 Number missing: 15, Reason: Adverse events: 8. Lack of efficacy: 2. Other: 5 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).; Group 2 Number missing: 9, Reason: Adverse events: 1. Lack of efficacy: 6. Other: 2 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).

Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Gastroduodenal ulcers and bleeding at 6 weeks; Group 1: 0/115, Group 2: 0/52; Comments: There were no episodes of gastroduodenal ulcers or bleeding
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, duration of osteoarthritis, primary study joint, pre-study therapy, and mini-mental state examination score. Does not report baseline values for outcomes.; Group 1 Number missing: 15, Reason: Adverse events: 8. Lack of efficacy: 2. Other: 5 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).;

Group 2 Number missing: 9, Reason: Adverse events: 1. Lack of efficacy: 6. Other: 2 (other includes deviations from protocol, withdrawal of consent, and loss to follow up). Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 6 weeks; Group 1: 3/115, Group 2: 0/52; Comments: Nabumetone: Headache = 3. Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, duration of osteoarthritis, primary study joint, pre-study therapy, and mini-mental state examination score. Does not report baseline values for outcomes.; Group 1 Number missing: 15, Reason: Adverse events: 8. Lack of efficacy: 2. Other: 5 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).; Group 2 Number missing: 9, Reason: Adverse events: 1. Lack of efficacy: 6. Other: 2 (other includes deviations from protocol, withdrawal of consent, and loss to follow up).	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Tugwell 2004[192]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=622)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic primary osteoarthritis of the knee and a recent (within 3 months) radiographic examination showing "osteoarthritis"
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and nonpregnant women between 40 and 85 years old with symptomatic primary osteoarthritis of the knee and a recent (within 3 months) radiographic examination showing "osteoarthritis". After a washout period people had to have a WOMAC pain subscale total score of at least 125mm (flare of pain not required), a WOMAC physical function subscale total score of at least 425mm and a patient global assessment score of at least 25mm.
Exclusion criteria	Secondary arthritis related to syphilitic neuropathy, ochronosis, psoriasis, metabolic bone disease, or acute trauma; chondrocalcinosis with a history of pseudogout; fibromyalgia; previous major surgery to the knee or recommendation for knee replacement/reconstruction; recent intraarticular viscosupplementation; current or recent corticosteroid use (orally, intramuscularly, or topically); topical product use at the application site; history of sensitivity to any of the study drugs, acetylsalicylic acid or other NSAIDs; severe, uncontrolled cardiac, renal, hepatic or other systemic disease; documented recent gastroduodenal ulcer or GI bleeding; history of alcohol or drug abuse; lactation; concomitant skin diseases at the application site; clinically significant elevation of serum creatinine (≥176.8 micromol/L) or of aspartate aminotransferase or alanine aminotransferase (≥3 times the upper limit of normal); or involvement within the previous 30 days in another investigational drug trial
Recruitment/selection of patients	After a washout period of 3-10 days (of all NSAIDs, narcotic analgesics, acetaminophen and other prohibited medications/therapies) people had to have a WOMAC pain subscale total score of at least 125mm (flare of pain not required), a WOMAC physical function subscale total score of at least 425mm and a patient global assessment score of at least 25mm.

Age, gender and ethnicity	Age - Mean (SD): 63.5 (10.0). Gender (M:F): 266:356. Ethnicity: White = 585, Asian = 5, Black = 7, Hispanic = 1, Other = 24
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Total x-ray score (maximum score 27) (mean [SD]): 6.3 (3.7) Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=311) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. Topical diclofenac solution (1.5% (w/w) diclofenac solution in a patented carrier including 45.5% (w/w) dimethyl sulfoxide) plus oral placebo capsules - 50 drops of study solution around the affected knee (20 drops to the front, 10 drops to each side and the base) without massage 3 times daily with the placebo capsule 3 times daily.. Duration 12 weeks. Concurrent medication/care: Concomitant medications, including NSAID, acetylsalicylic acid, paracetamol and other analgesic medications were prohibited during the study, but people were allowed to continue stable acetylsalicylic acid therapy (up to 325 mg/day) for cardiovascular prophylactic purposes. Indirectness: No indirectness (n=311) Intervention 2: NSAIDs - Diclofenac. Placebo solution (2.3% (w/w) dimethyl sulfoxide) plus ora diclofenac 50mg capsules - 50 drops of placebo around the affected knee (20 drops to the front, 10 drops to each side and the base) without massage 3 times daily with the capsule 3 times daily.. Duration 12 weeks. Concurrent medication/care: Concomitant medications, including NSAID, acetylsalicylic acid, paracetamol and other analgesic medications were prohibited during the study, but people were allowed to continue stable acetylsalicylic acid therapy (up to 325 mg/day) for cardiovascular prophylactic purposes. Indirectness: No indirectness
Funding	Study funded by industry (Supported by Dimethaid Healthcare Ltd.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus DICLOFENAC

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -127 (SD 120); n=237, Group 2: mean -140 (SD 127); n=255; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline topical: 286 (88). Baseline oral: 286 (100).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height,

heart rate, blood pressure, X-ray score, WOMAC subscales and patient global assessment; Group 1 Number missing: 129, Reason: 129 discontinued: 64 due to adverse events, 28 due to lack of effect, 5 lost to follow up, 32 other; Group 2 Number missing: 116, Reason: 116 discontinued: 79 due to adverse events, 10 due to lack of effect, 5 lost to follow up, 22 other Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -380 (SD 396); n=237, Group 2: mean -451 (SD 431); n=255; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Baseline topical: 968 (291). Baseline oral: 975 (332). Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, X-ray score, WOMAC subscales and patient global assessment; Group 1 Number missing: 129, Reason: 129 discontinued: 64 due to adverse events, 28 due to lack of effect, 5 lost to follow up, 32 other; Group 2 Number missing: 116, Reason: 116 discontinued: 79 due to adverse events, 10 due to lack of effect, 5 lost to follow up, 22 other Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: All GI events (including: abdominal pain, constipation, diarrhoea, dyspepsia, flatulence, nausea, vomiting, melena) at 12 weeks; Group 1: 108/311, Group 2: 150/311; Comments: Diclofenac gel: Abdominal pain = 36, constipation = 25, diarrhoea = 27, dyspepsia = 48, flatulence = 30, nausea = 25, vomiting = 5, melena = 4. Diclofenac capsules: Abdominal pain = 67, constipation = 31, diarrhoea = 54, dyspepsia = 81, flatulence = 52, nausea = 41, vomiting = 7, melena = 7. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, weight, height, heart rate, blood pressure, X-ray score, WOMAC subscales and patient global assessment; Group 1 Number missing: 129, Reason: 129 discontinued: 64 due to adverse events, 28 due to lack of effect, 5 lost to follow up, 32 other; Group 2 Number missing: 116, Reason: 116 discontinued: 79 due to adverse events, 10 due to lack of effect, 5 lost to follow up, 22 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Uchio 2018[193] (Uchio 2018[194])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=354)
Countries and setting	Conducted in Japan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People needed to satisfy the American College of Rheumatology criteria for idiopathic knee osteoarthritis (knee pain, bone spurs detected on plain X-ray images, and at least one of age >50 years, morning stiffness resolving within 30 minutes, or crepitus)
Stratum	Knee:
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients aged 40-79 years if they experienced pain for at least 14 days/month during the 3 month period before visit 1 and had a BPI-Severity average pain score of at least 4. People needed to satisfy the American College of Rheumatology criteria for idiopathic knee osteoarthritis.
Exclusion criteria	Previous administration of duloxetine; inflammatory arthritis (e.g. rheumatoid arthritis) or autoimmune diseases (except Hashimoto's disease and type 1 diabetes); invasive treatment (e.g. joint lavage or intra-articular hyaluronic acid or steroid injections) in either knee within 1 month before Visit 1; arthroscopic surgery of the affected joint within 1 year before Visit 1 or a history of joint replacement or osteotomy; end-stage osteoarthritis (e.g., patients with loss of joint space or loss of articular cartilage) or patients scheduled to undergo surgery of the affected joint during the study; major depressive disorders based on the Mini International Neuropsychiatric Interview; and suicidal tendencies according to the Columbia-Suicide Severity Rating Scale
Recruitment/selection of patients	Multicenter. People taking NSAIDs and other therapeutic agents for knee osteoarthritis were washed out during the screening phase, their concomitant use was prohibited during the study
Age, gender and ethnicity	Age - Mean (SD): 66.0 (8.2). Gender (M:F): 79:274. Ethnicity: Japanese people
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren-Lawrence grade 1-4. Duration of symptoms (mean [SD]): 4.2 (4.3) years. Some results are extracted from clinicaltrials.gov, NCT02248480 (https://clinicaltrials.gov/ct2/show/results/NCT02248480?term=NCT02248480&draw=2&rank=1)
Indirectness of population	No indirectness
Interventions	(n=178) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 20mg for 1 week, 40mg for 1 week, 60mg for 12 weeks. Duration 14 weeks. Concurrent medication/care: Drugs with analgesic effect (e.g. NSAIDs) were

	permitted as rescue medication for up to 3 consecutive days and a cumulative total of 20 days. Indirectness: No indirectness (n=176) Intervention 2: Placebo. Matching placebo. Duration 14 weeks. Concurrent medication/care: Drugs with analgesic effect (e.g. NSAIDs) were permitted as rescue medication for up to 3 consecutive days and a cumulative total of 20 days. Indirectness: No indirectness
Funding	Study funded by industry (This work was supported by Shionogi & Co., Ltd., Eli Lilly Japan K.K., and Eli Lilly and Company.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DULOXETINE versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D at 14 weeks; Group 1: mean 0.12 (SD 0.14); n=177, Group 2: mean 0.07 (SD 0.14); n=176; EQ-5D 0-1 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Calculated standard deviation from this. Reported duloxetine: 0.12 (0.10 to 0.14). Reported placebo: 0.07 (0.05 to 0.09).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, BMI, duration of osteoarthritis, baseline values for BPI average pain (but not WOMAC or EQ-5D), use of NSAIDs, prior drug therapy and Kellgren-Lawrence grade; Group 1 Number missing: 17, Reason: 2 patient request, 11 adverse events, 4 lack of efficacy or worsening disease; Group 2 Number missing: 14, Reason: 1 ineligible, 4 patient request, 2 adverse events, 6 lack of efficacy or worsening disease, 1 other reason

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 14 weeks; Group 1: mean -3.99 (SD 2.79); n=177, Group 2: mean -2.43 (SD 2.79); n=176; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports least square means and standard error. Calculated standard deviation from this. Reported duloxetine: -3.99 (0.21). Reported placebo: -2.43 (0.21).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, BMI, duration of osteoarthritis, baseline values for BPI average pain (but not WOMAC or EQ-5D), use of NSAIDs, prior drug therapy and Kellgren-Lawrence grade; Group 1 Number missing: 17, Reason: 2 patient request, 11 adverse events, 4 lack of efficacy or worsening disease; Group 2 Number missing: 14, Reason: 1 ineligible, 4 patient request, 2 adverse events, 6 lack of efficacy or worsening disease, 1 other reason

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 14 weeks; Group 1: mean -11.77 (SD 8.91); n=177, Group 2: mean -7.07 (SD 8.76); n=176; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reported least squares mean and standard error. Reported duloxetine: -11.77 (0.67). Reported placebo: -7.07 (0.66).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, BMI, duration of

osteoarthritis, baseline values for BPI average pain (but not WOMAC or EQ-5D), use of NSAIDs, prior drug therapy and Kellgren-Lawrence grade; Group 1 Number missing: 17, Reason: 2 patient request, 11 adverse events, 4 lack of efficacy or worsening disease; Group 2 Number missing: 14, Reason: 1 ineligible, 4 patient request, 2 adverse events, 6 lack of efficacy or worsening disease, 1 other reason Protocol outcome 4: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Palpitations at 14 weeks; Group 1: 1/178, Group 2: 0/176 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, weight, height, BMI, duration of osteoarthritis, baseline values for BPI average pain (but not WOMAC or EQ-5D), use of NSAIDs, prior drug therapy and Kellgren-Lawrence grade; Group 1 Number missing: 17, Reason: 2 patient request, 11 adverse events, 4 lack of efficacy or worsening disease; Group 2 Number missing: 14, Reason: 1 ineligible, 4 patient request, 2 adverse events, 6 lack of efficacy or worsening disease, 1 other reason	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study (subsidiary papers)	Underwood 2008[196] (Underwood 2008[195])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=282)
Countries and setting	Conducted in United Kingdom; Setting: General practice
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged at least 50 years, have had troublesome pain in or around the knee on most days for at least a month as well as knee pain for more than 3 months in the preceding year; have consulted or been prescribed treatment by the general practitioner for knee pain in the preceding three years; have no current or planned knee replacements; and meet our safety criteria. Each participant was assessed by a general practitioner who confirmed they were willing to prescribe NSAIDs for this participant and recorded the physical components of the American College of Rheumatologists' clinical criteria for knee osteoarthritis.
Exclusion criteria	Peptic ulceration (past or current); indigestion on most days in the past three months; previous adverse reaction to NSAIDs; raised blood pressure no less than 155/95mmHg (mean of three readings at study entry); uncontrolled heart failure; serum creatinine concentration >140 micromol/L; abnormal liver function sufficient to contraindicate use of NSAIDs; taking anticoagulants or oral steroids; haemoglobin <124g/L for men or <118g/L for women; disseminated malignancy; request by general practitioner not to include potential participant for any other reason
Recruitment/selection of patients	Potential participants were asked not to use any topical or oral NSAIDs for one week before the assessment at study entry.
Age, gender and ethnicity	Age - Median (IQR): Oral = 60 (56-69). Topical: 60 (56-68).. Gender (M:F): 118:151. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=138) Intervention 1: NSAID gels (topical - local) - Ibuprofen gel. Topical ibuprofen gel. Duration 24 months. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=144) Intervention 2: NSAIDs - Ibuprofen. Oral ibuprofen up to a maximum of 1.2grams per day. Duration 24 months. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Equipment / drugs provided by industry (This study was commissioned by the NHS Health Technology Assessment Programme, project reference 01/09/02. Goldshield Pharmaceuticals supplied the starter packs of topical ibuprofen. MU had received speaker fees from Pfizer.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: IBUPROFEN GEL versus IBUPROFEN

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: SF-36 physical component score at 3 months; MD; -0.1 (95%CI -1.7 to 1.8) SF-36 physical component summary 0-100 Top=High is good outcome, Comments: Baseline topical: 39.2 (8.9). Baseline oral: 39.0 (9.7);
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 8, Reason: 130 people completed the questionnaire; Group 2 Number missing: 10, Reason: 134 people completed the questionnaire
- Actual outcome for Knee: SF-36 physical component score at 24 months; MD; -0.7 (95%CI -2.5 to 1.2) SF-36 physical component score 0-100 Top=High is good outcome, Comments: Baseline topical: 39.2 (8.9). Baseline oral: 39.0 (9.7);
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 55, Reason: 83 people completed the questionnaire; Group 2 Number missing: 61, Reason: 83 people completed the questionnaire
- Actual outcome for Knee: SF-36 mental component score at 3 months; MD; -1.2 (95%CI -3.3 to 0.9) SF-36 mental component score 0-100 Top=High is good outcome, Comments: Baseline topical: 53.7 (9.6). Baseline oral: 52.0 (10.2);
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 8, Reason: 130 people completed the questionnaire; Group 2 Number missing: 10, Reason: 134 people completed the questionnaire

- Actual outcome for Knee: SF-36 mental component score at 24 months; MD; -0.5 (95%CI -2.6 to 1.7) SF-36 mental component score 0-100 Top=High is good outcome, Comments: Baseline topical: 53.7 (9.6). Baseline oral: 52.0 (10.2);
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 55, Reason: 83 people completed the questionnaire; Group 2 Number missing: 61, Reason: 83 people completed the questionnaire

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 3 months; MD; -2 (95%CI -6 to 2) WOMAC pain subscale 0-100 Top=High is poor outcome, Comments: Baseline topical: 39 (19.3). Baseline oral: 39 (21.5).;
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 8, Reason: 130 people completed the questionnaire; Group 2 Number missing: 10, Reason: 134 people completed the questionnaire
- Actual outcome for Knee: WOMAC pain subscale at 24 months; MD; +5 (95%CI 0 to 9) WOMAC pain subscale 0-100 Top=High is poor outcome, Comments: Baseline topical: 39 (19.3). Baseline oral: 39 (21.5).;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 55, Reason: 83 people completed the questionnaire; Group 2 Number missing: 61, Reason: 83 people completed the questionnaire

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC difficulty subscale at 3 months; MD; -2 (95%CI -5 to 2) WOMAC physical function subscale 0-100 Top=High is poor outcome, Comments: Baseline topical: 37 (18.3). Baseline oral: 38 (23.1).;
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 8, Reason: 130 people completed the questionnaire; Group 2 Number missing: 10, Reason: 134 people completed the questionnaire
- Actual outcome for Knee: WOMAC difficulty subscale at 24 months; MD; +3 (95%CI -2 to 7) WOMAC physical function subscale 0-100 Top=High is poor outcome, Comments: Baseline topical: 37 (18.3). Baseline oral: 38 (23.1).;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 55, Reason: 83 people completed the questionnaire; Group 2 Number missing:

61, Reason: 83 people completed the questionnaire Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastrointestinal adverse events at 24 months; Group 1: 58/138, Group 2: 57/144 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Includes bleeding gastrointestinal events and indigestion - best fits the bleeding adverse events outcome; Baseline details: Reports age, occupational codes, people who met the clinical criteria for osteoarthritis, mean BmI, WOmAC subscales, EQ-5D, chronic pain guide, SF-36, indigestion, knee pain, NSAID usage, blood pressure, lung function tests and blood test values; Group 1 Number missing: 55, Reason: 83 people completed the questionnaire; Group 2 Number missing: 61, Reason: 83 people completed the questionnaire	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	VACT trial: Schnitzer 2005[174]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (pooled analysis of 2 trials, VACT1 and VACT2 - However, VACT1 is reported in Geba 2002 [78] so this report will only extract VACT2, which is reported separately) (n=1578)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee defined by clinical diagnosis using American Rheumatology Association clinical criteria, which were designated as the primary source of pain or disability in the lower extremity for at least 6 months
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and nonpregnant women with symptomatic osteoarthritis of the knee who were at least 40 years of age, with a functional class rating of I-III who were previous users of either a single prescription strength nonselective NSAID or COX-2 selective inhibitor, or paracetamol for control of osteoarthritis symptoms for at least 30 days prior to study entry, who met baseline pain criteria (<80 or a 0-100 WOMAC VAS score at visit 1)
Exclusion criteria	Concurrent medical or arthritis disease or abnormal laboratory results (values outside the normal reference range or determined by the investigator to be of clinical significance) that had the potential to confound or interfere with the efficacy evaluation or would contraindicate participation in the trial; people with a history of allergy to the study drugs; hypersensitivity to aspirin, any non-selective NSAID or sulfonamide-containing compounds, or who received an investigational drug within 30 days of screening
Recruitment/selection of patients	Flare trial. Multicenter (97 study sites in the United States).
Age, gender and ethnicity	Age - Mean (SD): 62.1 (10.34). Gender (M:F): 515:1063. Ethnicity: White = 1384, Hispanic American = 63, Black = 107, Other = 24
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III Duration of symptoms: Not stated. Flare design - Medication was discontinued after

	visit 1. Only people who had significant worsening of knee pain or related symptoms at the end of their allowed washout period (or prior to this) were allowed to continue the study except in prior paracetamol users, who were accepted to have milder disease and due to paracetamol being available as rescue therapy during the phase
Indirectness of population	No indirectness
Interventions	(n=269) Intervention 1: Paracetemol (oral) - Paracetemol. Paracetamol 1000mg four times a day (4000mg in total). Duration 6 weeks. Concurrent medication/care: No other rescue medication was permitted during the trial. Indirectness: No indirectness (n=523) Intervention 2: NSAIDs - Celecoxib. Celecoxib 200mg per day in one dose with three placebo tablets (to make up four doses per day, as with paracetamol). Duration 6 weeks. Concurrent medication/care: No other rescue medication was permitted during the trial. Indirectness: No indirectness (n=786) Intervention 3: NSAIDs - Other. Rofecoxib either 12.5mg per day or 25mg per day with matching placebo to make up 4 doses per day (as with paracetamol). Duration 6 weeks. Concurrent medication/care: No other rescue medication was permitted during the trial. Indirectness: No indirectness Comments: Rofecoxib is not licensed for use in the UK and so will not be included in the analysis. It is reported here for completeness.
Funding	Study funded by industry (Funding for the VACT studies provided by Merck & Co. Inc., West Point, PA, USA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 6 weeks; Group 1: mean -29.6 (SD 25.8); n=523, Group 2: mean -24.6 (SD 25.8); n=269; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and the p-value. Calculated SD from this. Reported p-value = ≤0.01. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, and prior drug type; Group 1 Number missing: 85, Reason: 13 adverse events, 42 lack of efficacy, 9 lost to follow up, 1 moved, 6 withdrew consent, 4 protocol deviations, 10 other; Group 2 Number missing: 64, Reason: 21 adverse events, 34 lack of efficacy, 2 lost to follow up, 1 moved, 2 withdrew consent, 1 protocol deviation, 3 other

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 6 weeks; Group 1: mean -25.7 (SD 23); n=523, Group 2: mean -20 (SD 23); n=269; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports change scores and the p-value. Calculated SD from this. Reported

p-value = ≤0.001. Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, and prior drug type; Group 1 Number missing: 85, Reason: 13 adverse events, 42 lack of efficacy, 9 lost to follow up, 1 moved, 6 withdrew consent, 4 protocol deviations, 10 other; Group 2 Number missing: 64, Reason: 21 adverse events, 34 lack of efficacy, 2 lost to follow up, 1 moved, 2 withdrew consent, 1 protocol deviation, 3 other Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Lower extremity oedema at 6 weeks; Group 1: 9/523, Group 2: 3/269 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, and prior drug type; Group 1 Number missing: 85, Reason: 13 adverse events, 42 lack of efficacy, 9 lost to follow up, 1 moved, 6 withdrew consent, 4 protocol deviations, 10 other; Group 2 Number missing: 64, Reason: 21 adverse events, 34 lack of efficacy, 2 lost to follow up, 1 moved, 2 withdrew consent, 1 protocol deviation, 3 other Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 22/523, Group 2: 12/269 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, race, and prior drug type; Group 1 Number missing: 85, Reason: 13 adverse events, 42 lack of efficacy, 9 lost to follow up, 1 moved, 6 withdrew consent, 4 protocol deviations, 10 other; Group 2 Number missing: 64, Reason: 21 adverse events, 34 lack of efficacy, 2 lost to follow up, 1 moved, 2 withdrew consent, 1 protocol deviation, 3 other	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Verkleij 2015[197]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=104)
Countries and setting	Conducted in Netherlands; Setting: Primary care
Line of therapy	1st line
Duration of study	Intervention + follow up: 4 weeks with intervention, 12 weeks follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fulfilling the clinical criteria of the American College of Rheumatology for knee osteoarthritis
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People age ≥45 years who consulted their GP with a new episode of non-traumatic knee pain, knee pain severity of 2 or more (on a 0-10 scale) and fulfilled the clinical criteria of the American College of Rheumatology for knee osteoarthritis
Exclusion criteria	If they had a contraindication for NSAIDs and/or paracetamol use (for example, myocardial infarction or stroke, or oral use of corticosteroids); an arthroplasty or osteotomy of the knee on the contralateral or unilateral side; if they already took NSAIDs or paracetamol at doses similar to or higher than the study dose; if they had had surgery; major trauma of the affected joint
Recruitment/selection of patients	The originally planned inclusion period (18 months) was extended by an additional 12 month to try and recruit more participants.
Age, gender and ethnicity	Age - Mean (SD): 64.0 (9.1). Gender (M:F): 39:65. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren and Lawrence grades 0-4. More people with grade 0-1 changes. Duration of symptoms: Majority greater than 3 months (some less than 3 weeks).. A pragmatic open-labelled randomised controlled trial
Indirectness of population	No indirectness
Interventions	(n=52) Intervention 1: NSAIDs - Diclofenac. Diclofenac (maximum daily intake of 3x50mg) for 2 weeks and, if required, an additional 1-2 weeks. Duration 4 weeks maximum. Concurrent medication/care: Usual care was provided by the GP to all patients (no definition of what this means). Indirectness: No indirectness (n=52) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol (maximum daily

	intake of 3x1000mg) for 2 weeks and, if required, an additional 1-2 weeks. Duration 4 weeks maximum. Concurrent medication/care: Usual care was provided by the GP to all patients (no definition of what this means). Indirectness: No indirectness
Funding	Academic or government funding (This study was supported by the NutsOhra Foundation in Amsterdam, the Netherlands, and partly funded by a programme grant of the Dutch Arthritis Foundation for its centre of excellent 'Osteoarthritis in primary care')

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC versus PARACETEMOL

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Knee: EQ-5D at 12 weeks; Group 1: mean 0.8 (SD 0.2); n=52, Group 2: mean 0.8 (SD 0.1); n=52; EQ-5D 0-1 Top=High is good outcome; Comments: Baseline diclofenac: 0.67 (0.25). Baseline paracetamol: 0.74 (0.16).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect); Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown

Protocol outcome 2: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: KOOS Pain at 12 weeks; Group 1: mean 37.4 (SD 21); n=52, Group 2: mean 34.8 (SD 19.4); n=52; KOOS pain subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac 50.7 (16.4). Baseline paracetamol: 47.9 (16.8).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect). Reports age, gender, BMI, duration of symptoms, Kellgren and Lawrence score, clinical ACR criteria, and baseline values for outcomes; Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown

Protocol outcome 3: Physical function at ≤3- or >3- months
- Actual outcome for Knee: KOOS Function at 12 weeks; Group 1: mean 31.4 (SD 20.2); n=52, Group 2: mean 28.4 (SD 19.5); n=52; KOOS function subscale 0-100 Top=High is poor outcome; Comments: Baseline diclofenac: 46.0 (17.7). Baseline paracetamol: 42.7 (18.4).
Risk of bias: All domain – Very high, Selection – High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect). Reports age, gender, BMI, duration of symptoms, Kellgren and Lawrence score, clinical ACR criteria, and baseline values for outcomes; Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown

Protocol outcome 4: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Knee: Gastrointestinal adverse events at 12 weeks; Group 1: 19/52, Group 2: 7/52 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect). Reports age, gender, BMI, duration of symptoms, Kellgren and Lawrence score, clinical ACR criteria, and baseline values for outcomes; Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown Protocol outcome 5: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Cardiovascular adverse events at 12 weeks; Group 1: 8/52, Group 2: 5/52 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect). Reports age, gender, BMI, duration of symptoms, Kellgren and Lawrence score, clinical ACR criteria, and baseline values for outcomes; Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown Protocol outcome 6: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Nervous system adverse events at 12 weeks; Group 1: 14/52, Group 2: 13/52 Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: EQ-5D value is different at baseline (7% difference between groups - given the treatments having the same final value, this has a significant effect). Reports age, gender, BMI, duration of symptoms, Kellgren and Lawrence score, clinical ACR criteria, and baseline values for outcomes; Group 1 Number missing: 4, Reason: 2 for personal circumstances, 1 loss of contact, 1 unknown; Group 2 Number missing: 6, Reason: 3 personal circumstances, 3 reason unknown	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Vojtassak 2011[199]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=334)
Countries and setting	Conducted in Czech Republic, Romania, Slovakia, United Kingdom; Setting: Outpatient follow up
Line of therapy	2nd line
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate-to-severe pain induced by osteoarthritis (as defined by the American College of Rheumatology) of the hip or knee
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female subjects, aged ≥40 years, with moderate-to-severe osteoarthritis pain defined as a mean weekly score of ≥5 on a scale of 0-10 for pain on average on the BPI scale. People must have suffered from chronic osteoarthritis pain in the target joint for more than 3 months, and their pain must not have been adequate controlled with daily analgesic (NSAID or paracetamol) treatment for the month before the beginning of the study.
Exclusion criteria	Regular treatment with an opioid in the 4 weeks before the screening visit (infrequent use of tramadol, codeine, tilidine, or dihydrocodeine for no more than 10 days in the 4 weeks before the screening visit was acceptable, but subjects were to stop any use of weak opioids at the screening visit); another type of continuous pain that stood out in comparison with osteoarthritis pain such as fibromyalgia, cervical radiculopathy, or chronic low back pain; any of the following 6 months before entering the study: major trauma to target joints, infection in target joints, radiologically apparent avascular necrosis in target joints, hyaluronan injections in the target joints; arthrodesis in the year or arthroscopy in the 2 months before entering study; planned treatment that could have altered the degree of pain withint he study period, subjects who were being treated with buprenorphine, albuphine, or pentazocine; corticosteroid injects in the 3 months before the start of the study
Recruitment/selection of patients	Multicenter trial carried out at 18 sites in 4 European countries
Age, gender and ethnicity	Age - Median (range): Hydromorphone: 65 (43-85). Placebo: 66 (40-87). Gender (M:F): 80:207. Ethnicity: All were Caucasian

Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed without imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).
Extra comments	Severity: Moderate-to-severe Duration of symptoms: Not stated explicitly. At least 3 months.
Indirectness of population	No indirectness
Interventions	(n=139) Intervention 1: Strong opioids (oral) - Hydromorphone. OROS hydromorphone hydrochloride 4mg once daily, which could be titrated up to a maximum daily dose of 32mg over 4 weeks before the study officially started with at least 3-4 days between dose increments.. Duration 12 weeks (after a 4 week titration phase). Concurrent medication/care: Paracetamol was allowed as rescue medication, provided that a subject did not exceed the total permitted daily dose (4grams per day until day 8 and then 2g per day for the remainder of the study). Indirectness: No indirectness (n=149) Intervention 2: Placebo. Placebo. Duration 12 weeks (after a 4 week titration phase). Concurrent medication/care: Paracetamol was allowed as rescue medication, provided that a subject did not exceed the total permitted daily dose (4grams per day until day 8 and then 2g per day for the remainder of the study). Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Janssen Cilag Medical Affairs, EMEA, a division of Janssen Pharmaceutica NV, Beerse, Belgium)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: HYDROMORPHONE versus PLACEBO

Protocol outcome 1: Quality of life at ≤3- or >3- months
- Actual outcome for Other: SF-36 pain subscale at 12 weeks; Group 1: mean 17.5 (SD 20.48); n=129, Group 2: mean 19.47 (SD 23.5); n=142; SF-36 pain subscale 0-100 Top=High is good outcome; Comments: Baseline hydromorphone: 27.7 (10.84). Baseline placebo: 27.8 (11.01).
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, BMI, most affected joint, and baseline outcome values; Group 1 Number missing: 55, Reason: 55 withdrew - 36 adverse events, 11 withdrew consent, 5 inefficacy, 1 lost to follow up, 1 investigator decision, 1 other reason; Group 2 Number missing: 22, Reason: 33 withdrew - 16 inefficacy, 8 withdrew consent, 7 adverse events, 2 investigator decision
- Actual outcome for Other: SF-36 physical functioning subscale at 12 weeks; Group 1: mean 13.59 (SD 19.71); n=132, Group 2: mean 14.72 (SD 24.08); n=144; SF-36 physical functioning subscale 0-100 Top=High is good outcome; Comments: Baseline hydromorphone: 25.0 (12.57). Baseline placebo: 27.8 (13.33).
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, BMI, most affected joint, and baseline outcome values; Group 1 Number missing: 55, Reason: 55 withdrew - 36 adverse events, 11 withdrew consent, 5 inefficacy, 1 lost to follow up, 1 investigator decision, 1 other reason; Group 2 Number missing: 22, Reason: 33 withdrew - 16 inefficacy, 8 withdrew consent, 7 adverse events, 2 investigator decision - Actual outcome for Other: SF-36 social functioning subscale at 12 weeks; Group 1: mean 7.29 (SD 23.42); n=132, Group 2: mean 9.55 (SD 24.11); n=144; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Baseline hydromorphone: 52.1 (20.82). Baseline placebo: 52.0 (21.28). Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, BMI, most affected joint, and baseline outcome values; Group 1 Number missing: 55, Reason: 55 withdrew - 36 adverse events, 11 withdrew consent, 5 inefficacy, 1 lost to follow up, 1 investigator decision, 1 other reason; Group 2 Number missing: 22, Reason: 33 withdrew - 16 inefficacy, 8 withdrew consent, 7 adverse events, 2 investigator decision Protocol outcome 2: Pain reduction at ≤3- or >3- months - Actual outcome for Other: WOMAC pain subscale at 12 weeks; Group 1: mean -3.74 (SD 4.49); n=131, Group 2: mean -3.86 (SD 4.52); n=143; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline hydromorphone: 11.8 (2.63). Baseline placebo: 11.5 (2.71). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, BMI, most affected joint, and baseline outcome values; Group 1 Number missing: 55, Reason: 55 withdrew - 36 adverse events, 11 withdrew consent, 5 inefficacy, 1 lost to follow up, 1 investigator decision, 1 other reason; Group 2 Number missing: 22, Reason: 33 withdrew - 16 inefficacy, 8 withdrew consent, 7 adverse events, 2 investigator decision Protocol outcome 3: Physical function at ≤3- or >3- months - Actual outcome for Other: WOMAC functional impairment subscale at 12 weeks; Group 1: mean -11.93 (SD 13.17); n=132, Group 2: mean -11.9 (SD 14.35); n=144; WOMAC functional impairement subscale 0-68 Top=High is poor outcome; Comments: Baseline hydromorphone: 41.2 (9.25). Baseline placebo: 39.8 (9.46). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, BMI, most affected joint, and baseline outcome values; Group 1 Number missing: 55, Reason: 55 withdrew - 36 adverse events, 11 withdrew consent, 5 inefficacy, 1 lost to follow up, 1 investigator decision, 1 other reason; Group 2 Number missing: 22, Reason: 33 withdrew - 16 inefficacy, 8 withdrew consent, 7 adverse events, 2 investigator decision	
Protocol outcomes not reported by the study	Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Wadsworth 2016[200]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=259)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographically confirmed primary osteoarthritis of the knee within 1 year prior to screening consistent with grade 2-3 Kellgren-Lawrence grading scale
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 40-85 years; radiographically confirmed primary osteoarthritis of the nee within 1 year prior to screening consistent with Grade 2 or 3 on the Kellgren-Lawrence grading scale; and were receiving stable analgesic therapy (i.e. at least 3 days per week for the previous month with an oral or topical NSAID or paracetamol). Enrolled people underwent a 3-14 day medication washout period prior to their first study treatment application and were required to demonstrate a 'moderate flare' (defined as a baseline minimum pain score of at least 5 using an 11 point numeric rating scale and at least a 2 unit worsening from the screening pain score) of pain in at least one knee following washout of the pain medication.
Exclusion criteria	Secondary osteoarthritis (defined as osteoarthritis resulting from a specific cause such as an injury or an effect of obesity, genetics, inactivity, or other disease) or symptomatic chondrocalcinosis (determined via radiographic diagnosis with symptoms being present) of the study knee; study drug contraindicated or known sensitivity to diclofenac or other NSAIDs; other severe medical condition; any abnormality that could confound interpretation of the safety results; uncontrolled diabetes; documented history of alcohol or drug abuse within 1 year prior to the screening visit; breastfeeding females; major surgery; previous damage to the study knee at any time; minor knee surgery (including cartilage repair, collateral ligament repair and arthroscopic debridement) to the study knee within 1 year of screening; corticosteroid injection into the primary study knee within 90 days of screening or into any other joint within 30 days of screening, or topical corticosteroid use on the study knee; intra-articular viscosupplementation in the primary study knee in the 6 months prior to screening; prior use of a stable opioid analgesic; any other painful or disabling conditions affecting the knee or leg, or disabling condition of the hands or skin disorder with

	current involvement of the hands or the knee(s); referral to an orthopaedic surgeon for consideration of, or been advised to have, knee replacement or knee reconstruction surgery; advanced osteoarthritis of the knee such that all cartilage was eroded (i.e. radiographic examination showed that the joint space was eliminated in both lateral and medial tibia/femoral compartment); regular headaches requiring paracetamol
Recruitment/selection of patients	Enrolled people underwent a 3-14 day medication washout period prior to their first study treatment application and were required to demonstrate a 'moderate flare' (defined as a baseline minimum pain score of at least 5 using an 11 point numeric rating scale and at least a 2 unit worsening from the screening pain score) of pain in at least one knee following washout of the pain medication.
Age, gender and ethnicity	Age - Mean (SD): 61.1 (9.2). Gender (M:F): 85:174. Ethnicity: White = 218, Black/African American = 40, Asian = 1. Out of all of these, 2 were Hispanic/Latino.
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Kellgren-Lawrence grade 2-3 Duration of symptoms: Not stated. NCT01119898
Indirectness of population	No indirectness
Interventions	(n=130) Intervention 1: NSAID gels (topical - local) - Diclofenac gel. 1.5% Diclofenac gel - 2mL applied per knee every 12 hours for 4 weeks. Duration 4 weeks. Concurrent medication/care: People were allowed paracetamol as rescue medication on an as-needed basis (up to 1950mg/day) during the treatment period, except for the 3 days prior to each clinic visit.. Indirectness: No indirectness (n=129) Intervention 2: Placebo. Placebo gel applied every 12 hours. Duration 4 weeks. Concurrent medication/care: People were allowed paracetamol as rescue medication on an as-needed basis (up to 1950mg/day) during the treatment period, except for the 3 days prior to each clinic visit.. Indirectness: No indirectness
Funding	Study funded by industry (The manuscript was supported by Horizon Pharma USA Inc. This study was supported by Mallinckrodt Pharmaceuticals)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC GEL versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: WOMAC pain subscale at 4 weeks; Group 1: mean -4.5 (SD 4.5); n=130, Group 2: mean -3.6 (SD 4.2); n=129; WOMAC pain	

subscale 0-20 Top=High is poor outcome; Comments: Baseline diclofenac gel: 12.4 (3.1). Baseline placebo: 12.6 (3.4). Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, height, weight, BMI, Kellgren-Lawrence grade, WOMAC subscale scores, PGA rating and NRS score; Group 1 Number missing: 8, Reason: 93.8% completed the study; Group 2 Number missing: 12, Reason: 90.7% completed the study Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC physical function subscale at 4 weeks; Group 1: mean -14.3 (SD 14.7); n=130, Group 2: mean -11.5 (SD 13.8); n=129; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline diclofenac gel: 42.9 (10.4). Baseline placebo: 43.3 (10.3). Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, ethnicity, height, weight, BMI, Kellgren-Lawrence grade, WOMAC subscale scores, PGA rating and NRS score; Group 1 Number missing: 8, Reason: 93.8% completed the study; Group 2 Number missing: 12, Reason: 90.7% completed the study	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Wang 2017[201]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=407)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 13 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: The American College of Rheumatology criteria for the diagnosis of osteoarthritis were used. For the knee: osteoarthritis disease criteria included knee pain, osteophytes (with radiographic evidence), and at least 1 of the following 3 conditions: age >50, morning stiffness <30min or crepitus. For the hip, osteoarthritis disease criteria included hip pain and at least 2 of the following 3 conditions: erythrocyte sedimentation rate <20mm/h, femoral or acetabular osteophytes (with radiographic evidence) or radiographic joint space narrowing.
Stratum	Knee:
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female outpatients of at least 40 years of age who met clinical and radiographic criteria for the diagnosis of osteoarthritis of the knee or hip, had pain for at least 14 days of each month for 3 months prior to study entry, and had a rating of at least 4 on the BPI 24-h Average Pain item (Question 3 or the BPI-modified short-form) during screening, prior to treatment.
Exclusion criteria	Any diagnosis of psychosis, bipolar disorder, schizoaffective disorder, current major depression disorder, anxiety disorders (excluding phobias), alcohol or eating disorders, or suicidal risk; people taking any excluded medications (analgesic agents including but not limited to NSAIDs, paracetamol and opioids) that could not be discontinued at the first study visit; people who were anticipated by the investigator to require use of excluded medications during the study.
Recruitment/selection of patients	People had to withdraw from previous medication, but there was no inclusion criteria based on response after this.
Age, gender and ethnicity	Age - Mean (SD): 60.5 (8.3). Gender (M:F): 96:311. Ethnicity: Chinese patients
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Knee or hip).

Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 8.0 (7.5) years
Indirectness of population	No indirectness
Interventions	(n=205) Intervention 1: Antidepressants (oral) - SNRIs. Duloxetine 60mg once a day (started at 30mg once a day for 1 week). . Duration 13 weeks. Concurrent medication/care: No additional information (paracetamol, NSAIDs and opioids were not allowed). Indirectness: No indirectness Comments: The study is followed up by a 13 week open label treatment phase which is not included in this analysis (n=202) Intervention 2: Placebo. Matching placebo. Duration 13 weeks. Concurrent medication/care: No additional information (paracetamol, NSAIDs and opioids were not allowed). Indirectness: No indirectness Comments: The study is followed up by a 13 week open label treatment phase which is not included in this analysis
Funding	Study funded by industry (This study was funded by Eli Lilly and Company and/or one of its subsidiaries)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SNRIS versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain subscale at 13 weeks; Group 1: mean -3.03 (SD 2.84); n=184, Group 2: mean -2.32 (SD 2.82); n=182; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports least square mean change and 95% confidence intervals. Standard deviation calculated from this. Reported duloxetine: -3.03 (-3.44, -2.62). Reported placebo: -2.32 (-2.73, -1.91). Baseline duloxetine: 7.82 (2.96). Baseline placebo: 7.67 (2.99).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of osteoarthritis diagnosis, location of osteoarthritis, duration of osteoarthritis pain; Group 1 Number missing: 39, Reason: 20 adverse events, 16 withdrawal by subject, 2 lack of efficacy, 1 physician decision; Group 2 Number missing: 26, Reason: 10 adverse events, 12 withdrawal by subject, 3 lack of efficacy, 1 physician decision

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Other: WOMAC physical function subscale at 13 weeks; Group 1: mean -9.64 (SD 9.2); n=184, Group 2: mean -7.28 (SD 9.02); n=182; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Reports least square mean change and 95% confidence intervals. Standard deviation calculated from this. Reported duloxetine: -9.64 (-10.97, -8.31). Reported placebo: -7.28 (-8.59, -5.97). Baseline duloxetine: 26.73 (10.36). Baseline placebo: 27.17 (10.27).
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, duration of osteoarthritis

diagnosis, location of osteoarthritis, duration of osteoarthritis pain; Group 1 Number missing: 39, Reason: 20 adverse events, 16 withdrawal by subject, 2 lack of efficacy, 1 physician decision; Group 2 Number missing: 26, Reason: 10 adverse events, 12 withdrawal by subject, 3 lack of efficacy, 1 physician decision	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Wanka 1964[202]
Study type	RCT (Patient randomised; Crossover: 0 days)
Number of studies (number of participants)	1 (n=18)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (4 weeks for each intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip verified with clinical and radiological features
Stratum	Hip
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the hip verified with clinical and radiographic features
Exclusion criteria	None stated
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (range): 65 (50-78). Gender (M:F): 8:10. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Low comorbidity score (8 of 18 had comorbidities. The majority with 1, 2 people had 2.). 4. Site of osteoarthritis (for systemic treatments only): Hip
Extra comments	Severity: Radiographic severity 1-4 Duration of symptoms (mean [range]): 6 (1-14) years
Indirectness of population	No indirectness
Interventions	(n=18) Intervention 1: NSAIDs - Indomethacin. Indomethacin 25mg three times a day for one week, 50mg twice a day for one week, then 50mg three times a day for two weeks. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness (n=18) Intervention 2: Placebo. Matching placebo. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Equipment / drugs provided by industry (Dr R. Hodgkinson or Merck Sharp and Dohme Ltd., Research Division, supplied indomethacin and placebo tablets and made arrangements for the randomisation of the initial treatment)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: INDOMETHACIN versus PLACEBO Protocol outcome 1: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Hip: Palpitations at 4 weeks; Group 1: 1/18, Group 2: 0/18 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Doesn't report which people were receiving which treatment first. However, is a crossover study so populations should be comparable.; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Wasserman 1984[203]
Study type	RCT (Patient randomised; Crossover: 1 week before the study, 1 week after the study, 0 days in the middle (on placebo during this period for 2 weeks))
Number of studies (number of participants)	1 (n=14)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (1 week washout at the start and the end, 2 weeks of carprofen, 2 weeks of placebo and 2 weeks of indomethacin)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the hip, knee or both as judged by clinical criteria (pain on movement or tenderness) and radiological evidence of disease, such as osteophyte formation, marginal lipping, narrowing of the joint space, osteosclerosis of the joint margin, or bone cysts.
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Ambulatory outpatients if they had active symptomatic osteoarthritis of the hip knee or both. Osteoarthritis was considered active if the person had pain (or required analgesics) for at least 3 days per week for the preceding 6 months and required medication for osteoarthritis during the previous month.
Exclusion criteria	Any other joint disease in the past year (except arthritis due to trauma); a history of operation on any joint used for evaluation; osteoarthritis which was aggravated by trauma within the past 2 months; were hypersensitive to indomethacin, other NSAIDs, or propoxyphene; had any disease which precluded safety analysis; a history of active peptic ulcer or other gastrointestinal disorders; pregnant or nursing; any number of diseases including alcohol or drug abuse, diabetes, cardiopulmonary disease or epilepsy.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Range: 45 to 78 years. Gender (M:F): 4:10. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee (While it states knee, hip or both - it reports the majority had knee osteoarthritis).
Extra comments	Severity: Not stated Duration of symptoms: Not stated explicitly. At least 6 months.
Indirectness of population	No indirectness

Interventions	(n=14) Intervention 1: NSAIDs - Indomethacin. Indomethacin 25mg capsule with 3 placebo tablets once a day. Duration 2 weeks. Concurrent medication/care: Supplemental analgesics such as dextropropoxyphene or codeine, and supportive care involving heat, hydrotherapy and physiotherapy were allowed. The use of intra-articular injections of any experimental drug for one month prior to the study was not permitted for eight weeks prior to the study.. Indirectness: No indirectness (n=14) Intervention 2: NSAIDs - Other. Carprofen three 25mg tablets and a placebo capsule once per day. Duration 2 weeks. Concurrent medication/care: Supplemental analgesics such as dextropropoxyphene or codeine, and supportive care involving heat, hydrotherapy and physiotherapy were allowed. The use of intra-articular injections of any experimental drug for one month prior to the study was not permitted for eight weeks prior to the study.. Indirectness: No indirectness Comments: Carprofen is not licensed for use in the UK and so was not included in this analysis as agreed in the protocol. It was reported here for completeness. (n=14) Intervention 3: Placebo. Three placebo tablets and one placebo capsule once per day. Duration 2 weeks. Concurrent medication/care: Supplemental analgesics such as dextropropoxyphene or codeine, and supportive care involving heat, hydrotherapy and physiotherapy were allowed. The use of intra-articular injections of any experimental drug for one month prior to the study was not permitted for eight weeks prior to the study.. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: INDOMETHACIN versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Nausea and belching at 2 weeks; Group 1: 0/14, Group 2: 0/14; Comments: Indomethacin: Nausea = 0, belching = 0. Placebo: Nausea = 0, belching = 0.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values for age and gender; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 missing (no details)

Protocol outcome 2: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Other: Central nervous system adverse events (including headache, numbness, insomnia) at 2 weeks; Group 1: 1/14, Group 2: 2/13; Comments: Indomethacin = 1 headache. Placebo 1 headache, 1 numbness, upper extremities

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low, Comments - Unclear if randomised; Indirectness of outcome: No indirectness ; Baseline details: Reports overall values for age and gender; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 missing (no details)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Wiesenhutter 2005[204]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=528)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A clinical and radiographic diagnosis of osteoarthritis of the knee (tibiofemoral joint) or hip for at least the previous 6 months as well as people with newly diagnosed clinical symptoms of osteoarthritis in the study joint
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Otherwise healthy male and females aged 40 years or older who had a clinical and radiographic diagnosis of osteoarthritis of the knee (tibiofemoral joint) or hip for at least the previous 6 months as well as people with newly diagnosed clinical symptoms of osteoarthritis in the study joint that met American Rheumatism Association functional class I-III criteria for at least the preceding 6 months. The primary source of pain for each person was in the lower extremity. In people in whom both knees and/or hips were affected, the most painful joint was selected for study evaluation. Women of childbearing potential were determined to be in a nongravid state with use of a serum beta-human chorionic gonadotropin measurements and instructed to use contraceptive measures during the study. People who were regular NSAID users (at least 25 of the last 30 days preceding enrollment) were required to have experienced a positive therapeutic benefit for their osteoarthritis of the hip or knee after NSAID therapy. Prior NSAID users were required to have a prestudy score of less than 80mm on an 100mm visual analogue scale for patient assessment of walking on a flat surface. After cessation of NSAID therapy during an NSAID-specific "washout" period, these people were required to experience a flare of osteoarthritis pain. A flare was classified as sufficient if the minimum patient-reported pain score was 40mm while the patient walked on a flat surface, was at least 15mm greater than at the prestudy visit and had worsened by at least 1 unit (0-5 point Likert scale) in Investigator Global Assessment of Disease Status. People who were daily users of paracetamol (1.2-4grams) for at least 25 of the last 30 days preceding enrollment and had used no NSAIDs for the treatment of osteoarthritis were required to have minimum scores of 40mm for patient reported pain while walking on a flat surface and Patient Global

	Assessment of Disease Status and IGADS of fair, poor, or very poor. Because paracetamol acts only as an analgesic without the associated anti-inflammatory activity of etoricoxib and ibuprofen, the decision was made to limit the number of paracetamol users enrolled at each study site to 20%.
Exclusion criteria	Any past or current medical conditions such as joint injuries or rheumatologic, autoimmune or musculoskeletal disorders that could confound or interfere with efficacy evaluations. Use of intra-articular corticosteroids or hyaluronic acid injections to the study knee within the past 3 months, immunosuppressants within the past 3 months, corticosteroids by any systemic route, and hyaluronic injections or intra-articular corticosteroids for any joint in the past month were not permitted. People were excluded if they were required to take any antiplatelet agent other than aspirin (up to 100mg/d).
Recruitment/selection of patients	Flare criteria were required for previous NSAID users. These were adapted for paracetamol users with them limiting the number of people who were purely paracetamol users before the trial.
Age, gender and ethnicity	Age - Mean (SD): 62.0 (9.9). Gender (M:F): 156:372. Ethnicity: Asian = 3, Black = 29, White = 468, Other = 28
Further population details	1. Age: Mixed (Based on range, 40-89). 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: American Rheumatology Association functional class I-III, median class II Duration of symptoms (mean [SD]): 7.8 (7.9) years
Indirectness of population	No indirectness
Interventions	(n=424) Intervention 1: NSAIDs - Ibuprofen. Ibuprofen 800mg three times a day and matching placebo once a day or Etoricoxib 30mg once daily and matching placebo three times a day. Duration 12 weeks. Concurrent medication/care: Only paracetamol was permitted for rescue pain medication if needed. People taking medications for chronic conditions were required to continue taking stable doses 2 weeks before and throughout the 12 week study. People taking stable doses of glucosamine or chondroitin sulfate for at least 6 months before the study were allowed to enroll. Low dose aspirin (no more than 100mg/day) for cardioprophylaxis was permitted. Gastroprotective agents such as proton pump inhibitors, histamine-2 receptor blockers, sucralfate and misoprostol were allowed as necessary.. Indirectness: No indirectness Comments: The etoricoxib and ibuprofen groups were combined for class effect as agreed in the protocol

	(n=104) Intervention 2: Placebo. Matching placebo (matching for etoricoxib once a day, matching for ibuprofen three times a day). Duration 12 weeks. Concurrent medication/care: Only paracetamol was permitted for rescue pain medication if needed. People taking medications for chronic conditions were required to continue taking stable doses 2 weeks before and throughout the 12 week study. People taking stable doses of glucosamine or chondroitin sulfate for at least 6 months before the study were allowed to enroll. Low dose aspirin (no more than 100mg/day) for cardioprophylaxis was permitted. Gastroprotective agents such as proton pump inhibitors, histamine-2 receptor blockers, sucralfate and misoprostol were allowed as necessary.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Merck Research Laboratories, Rahway, NJ)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ETORICOXIB/IBUPROFEN versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: WOMAC pain walking on a flat surface at 12 weeks; Group 1: mean -30.7 (SD 24.9); n=424, Group 2: mean -20.3 (SD 23.8); n=104; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reports least square mean change and 95% confidence intervals. Converted into standard deviations. Reported etoricoxib (n=214): -30.92 (-34.23 to -27.61), reported ibuprofen (n=210): -30.48 (-33.87 to -27.09), reported placebo (n=104): -20.30 (-24.86 to -15.73). Baseline etoricoxib: 73.86 (15.45). Baseline ibuprofen: 73.16 (16.32). Baseline placebo: 72.83 (14.43). Calculated SD etoricoxib: 24.7. Calculated SD ibuprofen: 25.1. Calculated SD placebo: 23.8.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, mean height, mean weight, low dose aspirin use no more than 100mg per day; Group 1 Number missing: 104, Reason: Ibuprofen: 30 lack of efficacy, 22 clinical AEs, 2 laboratory adverse events, 10 other reasons. Etoricoxib: 25 lack of efficacy, 11 clinical AEs, 0 laboratory adverse events, 6 other reasons; Group 2 Number missing: 49, Reason: Placebo: 31 lack of efficacy, 6 clinical AEs, 0 laboratory adverse events, 12 other reasons

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Other: Discontinued study due to a digestive system or abdominal pain adverse event at 12 weeks; Group 1: 14/424, Group 2: 2/104; Comments: Including ... Etoricoxib: Heartburn = 2, nausea = 3. Ibuprofen: Heartburn = 5, nausea = 3. Placebo: Heartburn = 0, nausea = 4.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Discontinuation events; Baseline details: Reports sex, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, mean height, mean weight, low dose aspirin use no more than 100mg per day; Group 1 Number missing: 104, Reason: Ibuprofen: 30 lack of efficacy, 22 clinical AEs, 2 laboratory adverse events, 10 other reasons. Etoricoxib: 25 lack of efficacy, 11 clinical AEs, 0 laboratory adverse events, 6 other reasons; Group 2 Number missing: 49, Reason: Placebo: 31 lack

of efficacy, 6 clinical AEs, 0 laboratory adverse events, 12 other reasons.reasons. Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Other: Discontinuation due to oedema related adverse events and hypertension related adverse events at 12 weeks; Group 1: 8/424, Group 2: 1/104; Comments: Etoricoxib: Oedema related adverse events = 7, hypertension related adverse events = 4, adverse events of congestive heart failure, pulmonary oedema or cardiac failure = 0. Ibuprofen: Oedema related adverse events = 19, hypertension related adverse events = 8, adverse events of congestive heart failure, pulmonary oedema or cardiac failure = 0. Placebo: Oedema related adverse events = 0, hypertension related adverse events = 1, adverse events of congestive heart failure, pulmonary oedema or cardiac failure = 0 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Discontinuation events; Baseline details: Reports sex, race, age, primary osteoarthritis joint, mean duration of osteoarthritis, ARA functional class, mean height, mean weight, low dose aspirin use no more than 100mg per day; Group 1 Number missing: 104, Reason: Ibuprofen: 30 lack of efficacy, 22 clinical AEs, 2 laboratory adverse events, 10 other reasons. Etoricoxib: 25 lack of efficacy, 11 clinical AEs, 0 laboratory adverse events, 6 other reasons; Group 2 Number missing: 49, Reason: Placebo: 31 lack of efficacy, 6 clinical AEs, 0 laboratory adverse events, 12 other reasons.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Williams 1993[207]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=178)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: 1 of the following 2 sets of criteria for osteoarthritis: 1) a) knee pain while standing, walking, and/or on motion for at least 25 of the 30 days prior to initiation of treatment in the study; b) age ≥40 years; c) morning stiffness of <30 minutes in duration; d) knee crepitus on motion or 2) a) knee pain while standing, walking and/or on motion, for at least 25 of the 30 days prior to initiation of treatment in the study; b) osteophytes seen on knee radiograph
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People fulfilling the definition of knee osteoarthritis: 1 of the following 2 sets of criteria for osteoarthritis: 1) a) knee pain while standing, walking, and/or on motion for at least 25 of the 30 days prior to initiation of treatment in the study; b) age ≥40 years; c) morning stiffness of <30 minutes in duration; d) knee crepitus on motion or 2) a) knee pain while standing, walking and/or on motion, for at least 25 of the 30 days prior to initiation of treatment in the study; b) osteophytes seen on knee radiograph
Exclusion criteria	Osteoarthritis secondary to infectious arthritis, rheumatoid arthritis or another inflammatory arthropathy, psoriasis, haemochromatosis, ochronosis, calcium pyrophosphate deposition disease, Paget's disease, or neuropathy; presence of arserine bursitis, peptic ulcer disease, gastriti or chronic gastrointestinal disease, liver disease or elevated liver enzyme levels, significant renal or haematological disease; injury to the study knee within the preceding 3 months; anticipated inability of the person to comply with clinic visits throughout the duration of the study; previous allergy to naproxen or paracetamol; grade IV radiographic changes in the knee.
Recruitment/selection of patients	Multicenter trial
Age, gender and ethnicity	Age - Mean (SD): 59.5 (10.8). Gender (M:F): 44:134. Ethnicity: White = 107, Black = 64, Other = 7
Further population details	1. Age: Mixed (Range goes up to 85 years). 2. Diagnostic method: Mixed 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Mean radiographic severity (SD): 2.2 (0.94) Duration of symptoms: 79.8 (74.4) months.
Indirectness of population	No indirectness
Interventions	(n=90) Intervention 1: NSAIDs - Naproxen. Naproxen 375mg twice daily with a meal and 2 placebo tablets orally four times a day. Duration 2 years. Concurrent medication/care: Concurrent corticosteroid treatment was not permitted. Concomitant therapy with any experimental drug was not allowed.. Indirectness: No indirectness (n=88) Intervention 2: Paracetemol (oral) - Paracetemol. Paracetamol 325mg four times a day with two placebo tablets twice daily with meals. Duration 2 years. Concurrent medication/care: Concurrent corticosteroid treatment was not permitted. Concomitant therapy with any experimental drug was not allowed.. Indirectness: No indirectness
Funding	Academic or government funding (Supported by NIAMS (National Institute of Arthritis and Musculoskeletal and Skin Diseases) contract N01-AR-2264)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN versus PARACETEMOL

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Pain on motion (visual analogue scale) at 6 weeks; Group 1: mean -1.027 (SD 2.481); n=75, Group 2: mean -0.703 (SD 2.062); n=73; Visual analogue scale (pain on motion) 0-10 Top=High is poor outcome; Comments: Baseline naproxen: 5.289 (2.795). Baseline paracetamol: 4.753 (3.248).
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 15, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 15, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)
- Actual outcome for Knee: Pain on motion (visual analogue scale) at 2 years; Group 1: mean -2 (SD 3.2); n=35, Group 2: mean -1 (SD 2.9); n=27; Visual analogue scale (pain on motion) 0-10 Top=High is poor outcome; Comments: Baseline naproxen: 5.289 (2.795). Baseline paracetamol: 4.753 (3.248).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 55, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 61, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events (including diarrhoea) leading to withdrawal from the study at 2 years; Group 1: 17/90, Group 2: 6/88; Comments: Naproxen: Gastrointestinal total = 17, diarrhoea = 0. Paracetamol: Gastrointestinal total = 6, diarrhoea = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 55, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 61, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)

Protocol outcome 3: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months
- Actual outcome for Knee: Oedema and chest pain adverse events leading to withdrawal from the study at 2 years; Group 1: 2/90, Group 2: 3/88; Comments: Naproxen: Oedema = 1, chest pain = 1. Paracetamol: Oedema = 2, chest pain = 1.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 55, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 61, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)

Protocol outcome 4: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months
- Actual outcome for Knee: Hepatitis adverse events leading to withdrawal from the study at 2 years; Group 1: 1/90, Group 2: 0/88
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 55, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 61, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)

Protocol outcome 5: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months
- Actual outcome for Knee: Headache adverse events leading to withdrawal from the study at 2 years; Group 1: 0/90, Group 2: 1/88
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Withdrawal events; Baseline details: Reports age, sex, race, weight, duration of disease, radiographic severity, and baseline values of outcomes; Group 1 Number missing: 55, Reason: Not stated for 6 months. At 2 years, 21 adverse drug reactions, 14 lack of response, 20 other including noncompliance (10), concurrent illness (4), moved (5), protocol violation (1); Group 2 Number missing: 61, Reason: Not stated for 6 months. At 2 years, 16 adverse drug reactions, 19 lack of response, 26 other including noncompliance (18), concurrent illness (5), moved (3)

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Williams 2000[205]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=686)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee, diagnosed according to the American College of Rheumatology criteria (unclear if imaging was involved in the diagnosis)
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female adults with osteoarthritis of the knee, diagnosed according to the American College of Rheumatology criteria if they met the following entry criteria: osteoarthritis in a flare state; a functional capacity classification of I-III; a negative pregnancy test if female
Exclusion criteria	Any inflammatory arthritis other than osteoarthritis; gout; any joint trauma at the knee with osteoarthritis; had received any oral, intramuscular, intra-articular or soft tissue injections of corticosteroids within the 4 weeks before taking study medications; had taken any NSAID or analgesic agent within 4 hours (or at least five times the agent's half-life) before the arthritis assessments performed at the baseline visit (patients taking a stable dose of aspirin ≤325 mg/day for non-arthritis conditions were eligible); had any active gastrointestinal, renal, hepatic or coagulation disorders; had oesophageal or gastroduodenal ulceration within the previous 30 days; had any laboratory abnormalities within the previous 14 days considered by the investigator to be clinically significant
Recruitment/selection of patients	Conducted at 50 clinical sites in the United States
Age, gender and ethnicity	Age - Mean (SD): 62.8 (10.9). Gender (M:F): 230:456. Ethnicity: Asian = 1, Black = 56, Caucasian = 602, Hispanic = 22, Other = 5
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional capacity classification of I-III Duration of symptoms: 8.89 (8.33) years. Criteria for a flare state (after discontinuing current therapy) - if the person and physician's global assessment scores were "fair", "poor" or "very poor" and if three of the following four criteria were met: a person's

	assessment of arthritis pain-visual analogue scale measurement of at least 40mm; an increase from the screening visit of two or more points in the osteoarthritis severity index; an increase from the screening visit of one or more grades in the Patient's Global assessment; an increase of one or more grades from the screening visit in the Physician's Global Assessment. For people not receiving NSAIDs before the study, the criteria were altered: a person's assessment of arthritis pain-visual analogue scale measurement of at least 40 mm; an osteoarthritis severity index score of ≥7; a person's global assessment score of "poor" or "very poor"; a physician's global assessment score of "poor" or "very poor"
Indirectness of population	No indirectness
Interventions	(n=454) Intervention 1: NSAIDs - Celecoxib. Celecoxib 100mg twice a day or Celecoxib 200mg once a day (and a placebo tablet once a day). Duration 6 weeks. Concurrent medication/care: The use of NSAIDs, oral or injectable corticosteroids, analgesics, or anticoagulants was prohibited during the study. People taking aspirin (≤325mg/day for conditions other than arthritis for at least 30 days before the first dose of study medication) were permitted to continue the same dose regimen. Paracetamol up to 2g per day was allowed if taken for reasons other than relief of arthritis symptoms and for no more than 3 consecutive days. Paracetamol must have been avoided within 48 hours before arthritis assessments performed at any visit. Indirectness: No indirectness Comments: The two celecoxib groups were combined as agreed in the protocol (n=232) Intervention 2: Placebo. Matching placebo twice a day. Duration 6 weeks. Concurrent medication/care: The use of NSAIDs, oral or injectable corticosteroids, analgesics, or anticoagulants was prohibited during the study. People taking aspirin (≤325mg/day for conditions other than arthritis for at least 30 days before the first dose of study medication) were permitted to continue the same dose regimen. Paracetamol up to 2g per day was allowed if taken for reasons other than relief of arthritis symptoms and for no more than 3 consecutive days. Paracetamol must have been avoided within 48 hours before arthritis assessments performed at any visit. Indirectness: No indirectness
Funding	Study funded by industry (Supported by G.D. Searle & Co. and Pfizer Inc)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Knee: Total gastrointestinal adverse events, including: dyspepsia, diarrhoea, abdominal pain at 6 weeks; Group 1: 34/453, Group 2: 14/231; Comments: Celecoxib 100mg BD: Total = 17. Dyspepsia = 8, Diarrhoea = 3, Nausea = 2, Abdominal pain = 2. Celecoxib 200mg OD: Total = 17. Dyspepsia = 5, Diarrhoea = 5, Nausea = 2, Abdominal pain = 3. Placebo: Total = 14. Dyspepsia = 4, Diarrhoea = 3, Nausea = 5, abdominal pain = 3 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height and weight; Group 1 Number missing: 78, Reason: Treatment failure = 39, adverse event = 20, other = 19; Group 2 Number missing: 86, Reason: Treatment failure = 56, adverse event = 20, other = 10 Protocol outcome 2: Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months - Actual outcome for Knee: Peripheral oedema at 6 weeks; Group 1: 4/453, Group 2: 2/231; Comments: Celecoxib 100mg BD: 1, Celecoxib 200mg OD: 3 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height and weight; Group 1 Number missing: 78, Reason: Treatment failure = 39, adverse event = 20, other = 19; Group 2 Number missing: 86, Reason: Treatment failure = 56, adverse event = 20, other = 10 Protocol outcome 3: Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months - Actual outcome for Knee: Urinary tract infection at 6 weeks; Group 1: 4/453, Group 2: 2/231; Comments: Celecoxib 100mg BD: 1, Celecoxib 200mg OD: 3 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height and weight; Group 1 Number missing: 78, Reason: Treatment failure = 39, adverse event = 20, other = 19; Group 2 Number missing: 86, Reason: Treatment failure = 56, adverse event = 20, other = 10 Protocol outcome 4: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 35/453, Group 2: 18/231; Comments: Celecoxib 100mg BD: 17, Celecoxib 200mg OD: 18 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height and weight; Group 1 Number missing: 78, Reason: Treatment failure = 39, adverse event = 20, other = 19; Group 2 Number missing: 86, Reason: Treatment failure = 56, adverse event = 20, other = 10	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Pain reduction at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months

Study	Williams 2001[206]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=718)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a diagnosis of osteoarthritis of the knee as determined by the American college of Rheumatology clinical and radiographic criteria
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a diagnosis of osteoarthritis of the knee as determined by the American college of Rheumatology clinical and radiographic criteria of functional capacity classification I-III with a current flare of osteoarthritis of the knee.
Exclusion criteria	Inflammatory arthritis, gout, or joint trauma at the knee in additional to osteoarthritis; had received any oral, intramuscular, intra-articular, or soft-tissue injection of corticosteroids within the 4 weeks before taking study medication; had taken any NSAID or analgesic agent (with the exception of aspirin ≤325mg/day for conditions other than arthritis) within 48 hours of the baseline arthritis assessments; had an active GI, renal, hepatic or coagulation disorder; had oesophageal or gastroduodenal ulceration within the previous 30 days; had experienced NSAID hypersensitivity or any laboratory abnormalities considered by the investigator to be clinically significant within the previous 14 days; women of childbearing age if they were pregnant or were not using adequate contraception
Recruitment/selection of patients	Conducted over 98 clinical sites in the United States
Age, gender and ethnicity	Age - Mean (SD): 61.5 (11.9). Gender (M:F): 214:504. Ethnicity: 616 people were White, 77 people were Black, 22 were Hispanic, 1 was Asian, 2 were Other
Further population details	1. Age: <75 years 2. Diagnostic method: Diagnosed with imaging 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: Functional class I-III Duration of symptoms: 9.5 (8.5) years. Required people to have a flare before entering the study
Indirectness of population	No indirectness

Interventions	(n=474) Intervention 1: NSAIDs - Celecoxib. Celecoxib 100mg twice a day (n=243) or 200mg once a day (with a placebo capsule at the other time of the day) (n=231) . Duration 6 weeks. Concurrent medication/care: The use of NSAIDs, oral or injectable corticosteroids, analgesics, or anticoagulants was prohibited during the study. People taking aspirin ≤325mg/day for reasons other than arthritis, for ≥30 days before the first dose of study medication, were permitted to continue with the same dosing regimen. People were permitted to take up to 2 grams/day of paracetamol, for 3 consecutive days, for reasons other than relief of arthritis symptoms (however, it must not have been taken within 48 hours of any visits). Indirectness: No indirectness (n=244) Intervention 2: Placebo. Placebo twice a day. Duration 6 weeks. Concurrent medication/care: The use of NSAIDs, oral or injectable corticosteroids, analgesics, or anticoagulants was prohibited during the study. People taking aspirin ≤325mg/day for reasons other than arthritis, for ≥30 days before the first dose of study medication, were permitted to continue with the same dosing regimen. People were permitted to take up to 2 grams/day of paracetamol, for 3 consecutive days, for reasons other than relief of arthritis symptoms (however, it must not have been taken within 48 hours of any visits). Indirectness: No indirectness
Funding	Study funded by industry (This research was supported by Pharmacia Corporation Research and Development, Skokie, Illinois and Pfizer Inc, New York, New York)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: Visual analogue scale pain at 6 weeks; Group 1: mean 22.3 (SD 30.3); n=472, Group 2: mean 15 (SD 30.3); n=243; Visual analogue scale (pain) 0-100 Top=High is poor outcome; Comments: Reports means and a p-value. SD calculated from this. Reported celecoxib 100mg BD: 21.2. Reported celecoxib 200mg OD: 23.5. Reported placebo: 15.0. p value: <0.01. Calculated SE = 2.82. Baseline celecoxib 100mg BD: 67.5 (16.5). Baseline celecoxib 200mg OD: 65.2 (16.4). Baseline placebo: 68.2 (16.5).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height, weight, and baseline values for outcomes; Group 1 Number missing: 89, Reason: 194 in the 100mg BID arm and 191 in the 200mg OD arm completed the study with 472 being included in the ITT population; Group 2 Number missing: 79, Reason: 164 people completed the study with 243 being included in the ITT population

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months
- Actual outcome for Knee: Gastrointestinal adverse events at 6 weeks; Group 1: 83/472, Group 2: 34/243; Comments: This includes (but does not account for

all events) - Celecoxib 100mg BD: dyspepsia = 15, diarrhoea = 12. Celecoxib 200mg OD: dyspepsia = 17, diarrhoea = 7. Placebo: dyspepsia = 12, diarrhoea = 3. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height, weight, and baseline values for outcomes; Group 1 Number missing: 89, Reason: 194 in the 100mg BID arm and 191 in the 200mg OD arm completed the study with 472 being included in the ITT population; Group 2 Number missing: 79, Reason: 164 people completed the study with 243 being included in the ITT population Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Knee: Headache at 6 weeks; Group 1: 78/472, Group 2: 42/243; Comments: Celecoxib 100mg BD: 39. Celecoxib 200mg OD: 39. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, osteoarthritis duration, race, height, weight, and baseline values for outcomes; Group 1 Number missing: 89, Reason: 194 in the 100mg BID arm and 191 in the 200mg OD arm completed the study with 472 being included in the ITT population; Group 2 Number missing: 79, Reason: 164 people completed the study with 243 being included in the ITT population	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Wittenberg 2006[208]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=364)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 7 days
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate-to-severe symptomatic osteoarthritis of the knee, according to the American College of Rheumatology criteria. At screening people were required to be receiving NSAIDs/simple analgesics for their osteoarthritis and to have a pain intensity in the affected target joint of more than or equal to 40mm on a 100m visual analogue scale after activity (walking 20 paces on a flat surface).
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Male or female patients aged 50 years or older with moderate-to-severe symptomatic osteoarthritis of the knee, according to the American College of Rheumatology criteria. At screening people were required to be receiving NSAIDs/simple analgesics for their osteoarthritis and to have a pain intensity in the affected target joint of more than or equal to 40mm on a 100m visual analogue scale after activity (walking 20 paces on a flat surface).
Exclusion criteria	Secondary osteoarthritis; concomitant significant medical problems; a history of gastrointestinal bleeding, peptic ulcers or open knee surgery within 1 year of study entry; hypersensitivity to analgesics, antipyretics, NSAIDs or sulfonamides; people who had undergone observational arthroscopy, arthroscopic surgery, or lavage within the preceding 180 days; females who were pregnant, lactating, or of childbearing age and not using adequate means of contraception
Recruitment/selection of patients	After the initial screening period, people entered a 2-7 day washout period, during which any treatment NSAIDs or analgesics was discontinued (if applicable). After this people were required to have VAS actual pain intensity at baseline or more than or equal to 50mm for the most severely affected knee joint after activity (this was greater than that required at screening).
Age, gender and ethnicity	Age - Mean (SD): 65.0 (8.3). Gender (M:F): 153:211. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee

Extra comments	Severity: Moderate to severe Duration of symptoms (mean [SD]): 7.4 (7.1) years
Indirectness of population	No indirectness
Interventions	(n=145) Intervention 1: NSAIDs - Celecoxib. Celecoxib 200mg twice daily with placebo once daily. Duration 1 week. Concurrent medication/care: Concomitant treatment with histamine-2 receptor blockers, proton pump inhibitors, misoprostol, methotrexate, warfarin, analgesics (other than rescue medication) and systemic corticosteroids was not permitted during the study; neither was physiotherapy for the target joint. People who had received intra-articular corticosteroids in the study joint during the three months before the study were excluded. People receiving chondroitin and/or glucosamine were excluded. People were permitted to use rescue medication (paracetamol ≤3g/day) during the study, although use of rescue medication was prohibited from midnight before the baseline clinic visit.. Indirectness: No indirectness (n=144) Intervention 2: NSAIDs - Other. Lumiracoxib 400mg once a day and placebo tablets twice a day. Duration 1 week. Concurrent medication/care: Concomitant treatment with histamine-2 receptor blockers, proton pump inhibitors, misoprostol, methotrexate, warfarin, analgesics (other than rescue medication) and systemic corticosteroids was not permitted during the study; neither was physiotherapy for the target joint. People who had received intra-articular corticosteroids in the study joint during the three months before the study were excluded. People receiving chondroitin and/or glucosamine were excluded. People were permitted to use rescue medication (paracetamol ≤3g/day) during the study, although use of rescue medication was prohibited from midnight before the baseline clinic visit.. Indirectness: No indirectness Comments: Lumiracoxib is not licensed for use in the UK and so was not included in the final analysis as agreed in the protocol. It is reported here for completeness. (n=75) Intervention 3: Placebo. One placebo tablet (placebo lumircoxib) once daily and one placebo tablet (placebo celecoxib) twice daily. Duration 1 week. Concurrent medication/care: Concomitant treatment with histamine-2 receptor blockers, proton pump inhibitors, misoprostol, methotrexate, warfarin, analgesics (other than rescue medication) and systemic corticosteroids was not permitted during the study; neither was physiotherapy for the target joint. People who had received intra-articular corticosteroids in the study joint during the three months before the study were excluded. People receiving chondroitin and/or glucosamine were excluded. People were permitted to use rescue medication (paracetamol ≤3g/day) during the study, although use of rescue medication was prohibited from midnight before the baseline

	clinic visit.. Indirectness: No indirectness
Funding	Study funded by industry (This study was funded by Novartis Pharma AG. Preparation of the manuscript was supported by an educational grant from Novartis Pharma AG. Editorial support in preparing the manuscript was provided by Thomson ACUMED.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CELECOXIB versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Knee: WOMAC pain subscale at 1 week; Group 1: mean -4 (SD 3.3); n=145, Group 2: mean -2.7 (SD 3.2); n=75; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline celecoxib: 11.1 (2.9). Baseline placebo: 10.8 (2.6). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, VAS baseline values, and WOMAC baseline values; Group 1 Number missing: 1, Reason: 1 withdrew due to protocol violation; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Knee: WOMAC difficulty in performing daily activities subscale at 1 week; Group 1: mean -12.3 (SD 10.5); n=145, Group 2: mean -9 (SD 9.8); n=75; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline celecoxib: 38.8 (9.2). Baseline placebo: 38.3 (8.7). Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, VAS baseline values, and WOMAC baseline values; Group 1 Number missing: 1, Reason: 1 withdrew due to protocol violation; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Yocum 2000[209]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=774)
Countries and setting	Conducted in Unknown multicentre; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: At least a 3 month history of osteoarthritis of the knee or hip confirmed by x-ray and by clinical signs and symptoms, and pain on movement in the target joint
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were current NSAID users at least 40 years of age, at least a 3 month history of osteoarthritis of the knee or hip confirmed by X-ray and by clinical signs and symptoms, and pain on movement in the target joint
Exclusion criteria	Prior intolerance of any NSAIDs, analgesic or antipyretic; presence of aspirin hypersensitivity; any disease that, in the opinion of the investigator, could interfere with the evaluation of efficacy of safety; abnormal renal, haematologic or hepatic function; history of bleeding disorder or current therapy with an anticoagulant; recent (within 2 months) use of corticosteroids; treatment with intra-articular injections of hyaluronic acid in the prior 3 months; long-term use of gastrointestinal medications (histamine-2 receptor antagonists, misoprostol, proton pump inhibitors) that could not be discontinued prior to participation; history of narcotic and/or alcohol abuse; history of gastrointestinal perforations and peptic ulcer bleeding within the previous 6 months to enrollment
Recruitment/selection of patients	An NSAID-free period of at least 3 days was initiated, during which demonstration of a flare was required. A flare was defined as worsening of disease activity from initial screening that met the following criteria: at least 1 grade deterioration in the investigator's global assessment of disease activity; an increase of 10mm or greater on a 100mm visual analogue scale for patient's global assessment of disease activity; an increase greater than 35mm on an 100-mm visual analogue scale in the patient's overall assessment of pain.
Age, gender and ethnicity	Age - Mean (SD): 62.9 (10.3). Gender (M:F): 268:506. Ethnicity: White = 699, African American = 47, Other = 28

Further population details	1. Age: Mixed (The majority of people were under the age of 70 with an unclear proportion being over the age of 75). 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Hip or knee).
Extra comments	Severity: Not stated Duration of osteoarthritis (mean [SD]): 8.2 (8.3) years.
Indirectness of population	No indirectness
Interventions	(n=617) Intervention 1: NSAIDs - Diclofenac. Diclofenac 50mg twice daily and Meloxicam 3.75mg/day, 7.5mg/day or 15mg/day. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Comments: The meloxicam groups and diclofenac group were pooled together due to class effect as agreed in the protocol (n=157) Intervention 2: Placebo. Matching placebo. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness
Funding	Study funded by industry (This study was supported by a grant from Boehringer Ingelheim, Ridgefield, Conn, a manufacturer of meloxicam)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DICLOFENAC AND MELOXICAM versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: Pain at rest at 12 weeks; Group 1: mean -24.2 (SD 28.8); n=617, Group 2: mean -13.5 (SD 28.8); n=157; Pain at rest 0-100 Top=High is poor outcome; Comments: Reports least square means and p-values. Reported meloxicam 3.75mg (n=153): -21.3, p-value = <0.05. Reported meloxicam 7.5mg (n=153): -24.7, p-value = <0.001. Reported meloxicam 15mg (n=156): -25.8, p-value <0.001. Reported diclofenac (n=152): -25.1, p-value <0.001. Reported placebo (n=155): -13.5. Calculated SD meloxicam 3.75mg-placebo: 20.6. Calculated SD meloxicam 7.5mg-placebo: 29.6. Calculated SD meloxicam 15mg-placebo: 32.7. Calculated SD diclofenac-placebo: 30.6. Baseline meloxicam 3.75mg: 64.5. Baseline meloxicam 7.5mg: 63.8. Baseline meloxicam 15mg: 63.9. Baseline diclofenac: 63.6. Baseline placebo: 64.7.
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, race, age, target joint, osteoarthritis other than in target joint, duration of osteoarthritis, history of peptic ulcer bleeding, duration of prior NSAID use; Group 1 Number missing: 217, Reason: Meloxicam 3.75mg: Any adverse event = 13, lack of efficacy = 44, administrative/other = 8. Meloxicam 7.5mg: Any adverse event = 11, lack of efficacy = 26, administrative/other = 11. Meloxicam 15mg: Any adverse event = 13, lack of efficacy = 25, administrative/other = 5. Diclofenac: Any adverse event = 13, lack of efficacy = 16, administrative/other = 9.; Group 2 Number missing: 77, Reason: Placebo: Any adverse event = 6, lack of efficacy = 60, adminstrative/other = 9.

Protocol outcome 2: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months

- Actual outcome for Other: Any GI events (including: abdominal pain, constipation, diarrhoea, dyspepsia, flatulence, nausea, vomiting, other gastrointestinal events) at 12 weeks; Group 1: 131/617, Group 2: 27/157; Comments: Meloxicam 3.75mg: Abdominal pain = 2, constipation = 3, diarrhoea = 3, dyspepsia = 9, flatulence = 6, nausea = 9, vomiting = 0, other GI events = 4. Meloxicam 7.5mg: Abdominal pain = 3, constipation = 3, diarrhoea = 12, dyspepsia = 7, flatulence = 5, nausea = 6, vomiting = 2, other GI events = 6. Meloxicam 15mg: Abdominal pain = 4, constipation = 1, diarrhoea = 5, dyspepsia = 7, flatulence = 5, nausea = 6, vomiting = 2, other GI events = 9. Diclofenac: Abdominal pain = 2, constipation = 6, diarrhoea = 14, dyspepsia = 10, flatulence = 6, nausea = 11, vomiting = 4, other GI events = 11. Placebo: Abdominal pain = 4, constipation = 3, diarrhoea = 6, dyspepsia = 7, flatulence = 7, nausea = 5, vomiting = 3, other GI events = 4 Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, race, age, target joint, osteoarthritis other than in target joint, duration of osteoarthritis, history of peptic ulcer bleeding, duration of prior NSAID use; Group 1 Number missing: 217, Reason: Meloxicam 3.75mg: Any adverse event = 13, lack of efficacy = 44, administrative/other = 8. Meloxicam 7.5mg: Any adverse event = 11, lack of efficacy = 26, administrative/other = 11. Meloxicam 15mg: Any adverse event = 13, lack of efficacy = 25, administrative/other = 5. Diclofenac: Any adverse event = 13, lack of efficacy = 16, administrative/other = 9.; Group 2 Number missing: 77, Reason: Placebo: Any adverse event = 6, lack of efficacy = 60, adminstrative/other = 9. Protocol outcome 3: Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months - Actual outcome for Other: Headache at 12 weeks; Group 1: 47/617, Group 2: 16/157; Comments: Meloxicam 3.75mg: 13. Meloxicam 7.5mg: 12. Meloxicam 15mg: 13. Diclofenac: 9. Placebo: 16. Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, race, age, target joint, osteoarthritis other than in target joint, duration of osteoarthritis, history of peptic ulcer bleeding, duration of prior NSAID use; Group 1 Number missing: 217, Reason: Meloxicam 3.75mg: Any adverse event = 13, lack of efficacy = 44, administrative/other = 8. Meloxicam 7.5mg: Any adverse event = 11, lack of efficacy = 26, administrative/other = 11. Meloxicam 15mg: Any adverse event = 13, lack of efficacy = 25, administrative/other = 5. Diclofenac: Any adverse event = 13, lack of efficacy = 16, administrative/other = 9.; Group 2 Number missing: 77, Reason: Placebo: Any adverse event = 6, lack of efficacy = 60, adminstrative/other = 9.	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months

Study	Zautra 2005[210]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=104)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with osteoarthritis defined by the guidelines of the American College of Rheumatology
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed with osteoarthritis as defined by the guidelines of the American College of Rheumatology with moderate to severe pain in the most affected joint or site as defined by the following: 1) pain for at least 1 month before day 0 or after the person had discontinued their as-needed opioid; 2) pain during the week before day 0 with an average score of ≥5 (≥3 if receiving as needed opioids) on a numeric rating scale where 0 = no pain and 10 = pain as bad as you can imagine.
Exclusion criteria	If they were taking opioids at an average daily dose of >60mg of oxycodone equivalent during the month prior to the study; were allergic to opioids; were scheduled to have surgery during the study period; had unstable coexisting disease or active severe organ dysfunction; had active cancer; were pregnant or breast-feeding; or had a prior or present history of substance abuse; people who had intra-articular or intramuscular steroid injections involving the joint under evaluation within 6 weeks prior to baseline or during the study
Recruitment/selection of patients	9 participating clinics
Age, gender and ethnicity	Age - Mean (SD): 63.3 (11.6). Gender (M:F): 28:76. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Type of osteoarthritis not specified).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=56) Intervention 1: Strong opioids (oral) - Oxycodone. Controlled release oxycodone 10mg every 12 hours uptitrated to a maximum final dose of 12 tablets/day.

	Duration 2 weeks. Concurrent medication/care: Stable regimens of paracetamol, NSAIDs or oral steroids were allowed, but rescue medication was not. Indirectness: No indirectness (n=51) Intervention 2: Placebo. Matching placebo twice a day. Duration 2 weeks. Concurrent medication/care: Stable regimens of paracetamol, NSAIDs or oral steroids were allowed, but rescue medication was not. Indirectness: No indirectness
Funding	Study funded by industry (Supported in part by Purdue Pharma LP, Stamford, Connecticut)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYCODONE versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Other: 24 hour pain (visual analogue scale) at 2 weeks; MD; -1.45 (SE: 0.46) Visual analogue scale 0-10 Top=High is poor outcome, Comments: Baseline oxycodone: 6.61 (1.66). Baseline placebo: 6.81 (1.45).;
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Reports age, gender, ethnicity, emotionality, interpersonal sensitivity, 24-hour pain, positive affect, negative affect, coping efficacy, helplessness, passive coping, active coping; Group 1 Number missing: 1, Reason: Withdrew before the first posttest (no reason given); Group 2 Number missing: 2, Reason: Withdrew before the first posttest (no reason given)

Protocol outcome 2: Psychological distress at ≤3- or >3- months
- Actual outcome for Other: Negative affect at 2 weeks; MD; -0.20 (SE : 0.14) Negative affect scale (subscale of the positive affect negative affect scale) 0-10 Top=High is poor outcome, Comments: Baseline oxycodone: 2.31 (0.81). Baseline placebo: 2.159 (0.68).;
Risk of bias: All domain – High, Selection – High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Not a standard scale approved by the committee prior to the review; Baseline details: Reports age, gender, ethnicity, emotionality, interpersonal sensitivity, 24-hour pain, positive affect, negative affect, coping efficacy, helplessness, passive coping, active coping; Group 1 Number missing: 1, Reason: Withdrew before the first posttest (no reason given); Group 2 Number missing: 2, Reason: Withdrew before the first posttest (no reason given)

Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Physical function at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Zenk 2002[211]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Physician-diagnosed osteoarthritis and daily joint pain, stiffness and immobility
Stratum	Other
Subgroup analysis within study	Not applicable
Inclusion criteria	Outpatients aged at least 19 years with physician-diagnosed osteoarthritis and daily joint pain, stiffness and immobility.
Exclusion criteria	People who required continued prescription medication for arthritis; nonambulatory due to arthritis; an active major organ system disease; pregnant or lactating
Recruitment/selection of patients	No response related inclusion criteria
Age, gender and ethnicity	Age - Mean (SD): 58 (12). Gender (M:F): 5:30. Ethnicity: Not stated
Further population details	1. Age: <75 years 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Multisite (Unclear - not reported).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=14) Intervention 1: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Glucosamine 500mg three times a day. Duration 6 weeks. Concurrent medication/care: The use of approved rescue medications were permitted, including naproxen 220mg, ibuprofen 200mg, paracetamol 325mg, and acetylsalicylic acid 325mg.. Indirectness: No indirectness (n=14) Intervention 2: Placebo. Matching placebo. Duration 6 weeks. Concurrent medication/care: The use of approved rescue medications were permitted, including naproxen 220mg, ibuprofen 200mg, paracetamol 325mg, and acetylsalicylic acid 325mg.. Indirectness: No indirectness

	(n=14) Intervention 3: Glucosamine (licensed preparations only, oral) - Glucosamine (licensed preparations only). Milk protein concentrate 2000mg twice a day. Duration 6 weeks. Concurrent medication/care: The use of approved rescue medications were permitted, including naproxen 220mg, ibuprofen 200mg, paracetamol 325mg, and acetylsalicylic acid 325mg.. Indirectness: No indirectness Comments: Milk protein concentrate is not an intervention included in the protocol and so is not included in the analysis. It is reported here for completeness.
Funding	Study funded by industry (This trial was jointly sponsored by Stolle Milk Biologics, Inc, Cincinnati, Ohio, and Humanetics Corporation, Chanhassen, Minnesota)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: GLUCOSAMINE (LICENSED PREPARATIONS ONLY) versus PLACEBO Protocol outcome 1: Pain reduction at ≤3- or >3- months - Actual outcome for Other: WOMAC pain subscale at 6 weeks; Group 1: mean 16.2 (SD 25.8); n=13, Group 2: mean 0.5 (SD 15); n=10; WOMAC pain subscale 0-100 Top=High is good outcome; Comments: Baseline glucosamine: 58.5 (8.8). Baseline placebo: 76.0 (16.5). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC subscales are different at baseline (lower in the glucosamine group, higher in the placebo group). Reports age, sex, body weight, systolic blood pressure, diastolic blood pressure, heart rate and temperature.; Group 1 Number missing: 1, Reason: 1 withdrew after developing costochondritis; Group 2 Number missing: 4, Reason: 4 withdrew due to lack of compliance with the protocol (1), dysphagia (1), unhappiness with aspects of the study protocol (2) Protocol outcome 2: Physical function at ≤3- or >3- months - Actual outcome for Other: WOMAC activities subscale at 6 weeks; Group 1: mean 13.2 (SD 23.5); n=23, Group 2: mean 2.3 (SD 12); n=10; WOMAC activities subscale 0-100 Top=High is good outcome; Comments: Baseline glucosamine: 59.8 (12.4). Baseline placebo: 74.7 (14.4). Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC subscales are different at baseline (lower in the glucosamine group, higher in the placebo group). Reports age, sex, body weight, systolic blood pressure, diastolic blood pressure, heart rate and temperature.; Group 1 Number missing: 1, Reason: 1 withdrew after developing costochondritis; Group 2 Number missing: 4, Reason: 4 withdrew due to lack of compliance with the protocol (1), dysphagia (1), unhappiness with aspects of the study protocol (2)	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months

Study	Zhao 1999[212]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1004)
Countries and setting	Conducted in Canada, USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis and fulfilled American College of Rheumatology clinical criteria for a diagnosis of primary osteoarthritis of the knee present for 3 months or longer who were in a functional class I-III
Stratum	Knee
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women outpatients 18 years of age or older if they had symptomatic osteoarthritis. After a 2- to 7-day washout period of NSAIDs or other analgesics, symptomatic arthritis (flare) was confirmed at a baseline visit according to predefined criteria for changes in results of arthritis assessments performed at screening.
Exclusion criteria	People who did not have a flare after the washout period; people with serious concomitant gastrointestinal, renal, hepatic or coagulation disorders; malignancy; ulcerations; those diagnosed with any other inflammatory arthritis (e.g. rheumatoid arthritis); gout; or acute trauma of the knee
Recruitment/selection of patients	People had to have a flare of osteoarthritis after a washout period. A flare was defined as worsening of signs and symptoms of the disease after discontinuation of NSAIDs or other analgesics, or other criteria if the person was not receiving treatment.
Age, gender and ethnicity	Age - Mean (range): 62 (21-89). Gender (M:F): 282:722. Ethnicity: Not stated
Further population details	1. Age: Mixed 2. Diagnostic method: Not stated / Unclear 3. Multimorbidities: Not stated / Unclear 4. Site of osteoarthritis (for systemic treatments only): Knee
Extra comments	Severity: OA severity index (0-24) (mean [SD]): 15.5 (3.5) Duration of symptoms (mean [SD]): 9.4 (8.3) years.
Indirectness of population	No indirectness
Interventions	(n=873) Intervention 1: NSAIDs - Naproxen. Naproxen 500mg twice a day or Celecoxib 50mg, 100mg or 200mg twice a day for 12 weeks. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (≤325mg/day) and paracetamol (up to 2g/day for no longer than 3 consecutive days) were allowed as concomitant therapy

	except within 48 hours before assessments, during which no analgesics were allowed. Use of other NSAIDs, oral or injectable corticosteroids, and anticoagulants were prohibited.. Indirectness: No indirectness Comments: The naproxen and celecoxib groups were combined for class effect as agreed in the protocol (n=219) Intervention 2: Placebo. Placebo twice daily. Duration 12 weeks. Concurrent medication/care: Low dose aspirin (≤325mg/day) and paracetamol (up to 2g/day for no longer than 3 consecutive days) were allowed as concomitant therapy except within 48 hours before assessments, during which no analgesics were allowed. Use of other NSAIDs, oral or injectable corticosteroids, and anticoagulants were prohibited.. Indirectness: No indirectness
Funding	Study funded by industry (Supported by G.D. Searle & Company, Skokie, Illinois)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: NAPROXEN/CELECOXIB versus PLACEBO

Protocol outcome 1: Pain reduction at ≤3- or >3- months
- Actual outcome for Knee: WOMAC pain subscale at 12 weeks; Group 1: mean -2.6 (SD 4.9); n=873, Group 2: mean -1.2 (SD 3.8); n=219; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports least square means and a p-value versus placebo. SD calculated from this. Celecoxib 50mg: -2.0, Celecoxib 100mg: -3.1, Celecoxib 200mg: -2.7. Naproxen: -2.4. Placebo: -1.2. Calculated SE Celecoxib: 0.42. Calculated SD Celecoxib: 5.2. Calculated SE Naproxen: 0.36. Calculated SD Naproxen: 3.8. Baseline celecoxib 50mg: 10.7 (3.2). Baseline celecoxib 100mg: 10.5 (3.4). Baseline celecoxib 200mg: 10.7 (3.4). Baseline naproxen: 11.0 (3.0). Baseline placebo: 10.8 (3.4).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of disease, weight, BMI, severity index and baseline values of outcomes; Group 1 Number missing: 321, Reason: Celecoxib 50mg: 1 lost to follow up, 1 preexisting violation, 4 protocol noncompliance, 61 treatment failures, 18 adverse events. Celecoxib 100mg: 3 lost to follow up, 7 protocol noncompliance, 40 treatment failures, 31 adverse events. Celecoxib 200mg: 1 lost to follow up, 2 protocol noncompliance, 49 treatment failure, 21 adverse events. Naproxen: 3 lost to follow up, 1 preexisting violation, 8 protocol noncompliance, 52 treatment failure, 18 adverse events; Group 2 Number missing: 113, Reason: 3 lost to follow up, 3 preexisting violations, 12 protocol noncompliance, 79 treatment failure, 16 adverse events

Protocol outcome 2: Physical function at ≤3- or >3- months
- Actual outcome for Knee: WOMAC physical function subscale at 12 weeks; Group 1: mean -8 (SD 15.4); n=873, Group 2: mean -3.9 (SD 16.3); n=219; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports least square means and a p-value versus placebo. SD calculated from this. Celecoxib 50mg: -6.8, Celecoxib 100mg: -9.5, Celecoxib 200mg: -8.1. Naproxen: -7.8. Placebo: -3.9. Placebo: -3.9. Calculated SE Celecoxib: 1.3. Calculated SD Celecoxib: 16.3. Calculated SE Naproxen: 1.2. Calculated SD Naproxen: 12.3. Baseline celecoxib 50mg: 36.2 (10.8). Baseline celecoxib 100mg: 35.4 (11.8). Baseline celecoxib 200mg: 35.3 (12.3). Baseline naproxen: 36.6 (10.6). Baseline placebo: 36.0 (10.8).
Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of disease, weight, BMI, severity index and baseline values of outcomes; Group 1 Number missing: 321, Reason: Celecoxib 50mg: 1 lost to follow up, 1 preexisting violation, 4 protocol noncompliance, 61 treatment failures, 18 adverse events. Celecoxib 100mg: 3 lost to follow up, 7 protocol noncompliance, 40 treatment failures, 31 adverse events. Celecoxib 200mg: 1 lost to follow up, 2 protocol noncompliance, 49 treatment failure, 21 adverse events. Naproxen: 3 lost to follow up, 1 preexisting violation, 8 protocol noncompliance, 52 treatment failure, 18 adverse events; Group 2 Number missing: 113, Reason: 3 lost to follow up, 3 preexisting violations, 12 protocol noncompliance, 79 treatment failure, 16 adverse events Protocol outcome 3: Serious adverse event 1: Gastrointestinal adverse events at ≤3- or >3- months - Actual outcome for Knee: Gastric ulcer and gastrointestinal haemorrhage at 12 weeks; Group 1: 2/873, Group 2: 0/219; Comments: Naproxen: 1 gastrointestinal haemorrhage, 1 gastric ulcer Risk of bias: All domain – Very high, Selection – High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of disease, weight, BMI, severity index and baseline values of outcomes; Group 1 Number missing: 321, Reason: Celecoxib 50mg: 1 lost to follow up, 1 preexisting violation, 4 protocol noncompliance, 61 treatment failures, 18 adverse events. Celecoxib 100mg: 3 lost to follow up, 7 protocol noncompliance, 40 treatment failures, 31 adverse events. Celecoxib 200mg: 1 lost to follow up, 2 protocol noncompliance, 49 treatment failure, 21 adverse events. Naproxen: 3 lost to follow up, 1 preexisting violation, 8 protocol noncompliance, 52 treatment failure, 18 adverse events; Group 2 Number missing: 113, Reason: 3 lost to follow up, 3 preexisting violations, 12 protocol noncompliance, 79 treatment failure, 16 adverse events	
Protocol outcomes not reported by the study	Quality of life at ≤3- or >3- months; Psychological distress at ≤3- or >3- months; Osteoarthritis flare-ups at ≤3- or >3- months; Serious adverse event 2: Renal and hepatic adverse events at ≤3- or >3- months; Serious adverse event 2: Cardiovascular adverse events at ≤3- or >3- months; Serious adverse event 2: Central nervous system adverse events at ≤3- or >3- months