National Institute for Health and Clinical Excellence

Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures

Assessment Report

Commercial in Confidence stripped version for consultation

Produced by: School of Health and Related Research, University of Sheffield (ScHARR)

STRICTLY CONFIDENTIAL

Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Clinical Excellence

Title: Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: a systematic review and cost-effectiveness analysis

Produced by ScHARR, University of Sheffield

Authors Matt Stevenson, Professor of Health Technology Assessment, ScHARR
Tim Gomersall, Research Associate, ScHARR
Myfanwy Lloyd Jones, Senior Research Fellow, ScHARR
Andrew Rawdin, Cost-effectiveness Modeller, ScHARR
Monica Hernández, Research Fellow in Econometrics, ScHARR
Sofia Dias, Research Associate, Centre for Academic Primary Care, University of Bristol
David Wilson, Consultant Musculoskeletal Interventional Radiologist, Oxford University Hospitals
Angie Rees, Information Specialist, ScHARR

Correspondence to Matt Stevenson
Professor of Health Technology Assessment and Technical Director of the ScHARR Technology Assessment Group
ScHARR
University of Sheffield
Regent Court
30 Regent Street
Sheffield S1 4DA
Tel: 0114 222 0691
E-mail: M.D.Stevenson@sheffield.ac.uk

Date completed 8th August 2012

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 10/122.

Declared competing interests of the authors

David Wilson undertakes clinical practice in cement augmentation of the spine both in the NHS and private sector.

None for the remaining authors

Acknowledgements

Tony Ades, Professor of Public Health Science, Centre for Academic Primary Care, University of Bristol

We would like to thank the following individuals for their generous responses to our data requests: Anastasios Mpotsaris, Klinik für Radiologie, Neuroradiologie und interventionelle Therapie, Klinikum Vest, Recklinghausen, Rachelle Buchbinder, Director, Monash University Department of Epidemiology, and Professor in the Monash University Department of Epidemiology & Preventive Medicine; Margaret Staples, biostatistician, Monash University Department of Clinical Epidemiology, Cabrini Hospital; Douglas Wardlaw, Professor, Orthopaedic Department, Woodend Hospital,Aberdeen; and Jordi Blasco Andaluz, Neurointerventional Department, Centre Diagnòstic per la Imatge, Hospital Clinic, Barcelona. The authors wish to thank Gill Rooney, Programme Administrator, ScHARR for her help in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Stevenson M, Gomersall T, Lloyd-Jones M, Rawdin A, Hernández M, Dias S, Wilson D, & Rees A. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: a systematic review and cost-effectiveness analysis. Health Technol. Assess.

Contributions of authors

Matt Stevenson was the Assessment Group lead, undertook the cost-effectiveness review, and was involved in constructing the cost-effectiveness model and interpreting the results. Myfanwy Lloyd Jones and Tim Gomersall undertook the clinical effectiveness review. Andrew Rawdin was involved in constructing and running the cost-effectiveness model; Sofia Dias constructed the network meta-analysis. Monica Hernández critiqued the data indicating a mortality benefit for percutaneous vertebroplasty and percutaneous balloon kyphoplasty. David Wilson provided clinical expertise. Angie Rees performed the literature searches.

About ScHARR

The School of Health and Related Research (ScHARR) is one of the four Schools that comprise the Faculty of Medicine at the University of Sheffield. ScHARR brings together a wide range of medical and health-related disciplines including public health, general practice, mental health, epidemiology, health economics, management sciences, medical statistics, operational research and information science. It includes the Sheffield unit of the Trent Institute for Health Services Research, which is funded by NHS R&D to facilitate high-quality health services research and capacity development.

The ScHARR Technology Assessment Group (ScHARR-TAG) synthesises research on the clinical effectiveness and cost-effectiveness of healthcare interventions for the NIHR Health Technology Assessment Programme on behalf of a range of policy makers, including the National Institute for Health and Clinical Excellence (NICE). ScHARR-TAG is part of a wider collaboration of a number of units from other regions including Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Aberdeen Health Technology Assessment Group (Aberdeen HTA Group), University of Aberdeen; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; the NHS Centre for Reviews and Dissemination, University of York; Warwick Evidence, The University of Warwick; the BMJ Group and Kleijnen Systematic Reviews.

Word count:

107,406 (including appendices)

TABLE OF CONTENTS

1.	DEFINITION OF TERMS AND LIST OF ABBREVIATIONS	1
2.	EXECUTIVE SUMMARY	3
2.1	Background	3
2.2	Objectives	4
2.3	Methods	4
2.4	Results	5
2.5	Discussion	7
2.6	Conclusion	8
3	BACKGROUND	10
3.1	Description of health problem	10
3.2	Current service provision	22
3.3	Description of technologies under assessment	26
4.	DEFINITION OF THE DECISION PROBLEM	35
4.1	Decision problem	35
4.2	Overall aims and objectives of assessment	50
5.	ASSESSMENT OF CLINICAL EFFECTIVENESS	51
5.1	Methods for reviewing clinical effectiveness and safety	51
5.2	Results	57
6.	ASSESSMENT OF COST-EFFECTIVENESS	155
6.1	Systematic review of existing cost-effectiveness evidence	155
6.2	The Assessment Group model	177
7.	ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES	268
8.	DISCUSSION	270
8.1	Statement of principal findings	270
8.2	Strengths and limitations of the assessment	271
8.3	Uncertainties	272
8.4	Other relevant factors	273
9.	CONCLUSIONS	274
9.1	Implications for service provision	274
9.2	Suggested research priorities	274
10	APPENDICES	276
	Appendix 1: Literature Search Strategies	276
	Appendix 2: Data extraction form	278
	Appendix 3: Original quality assessment checklist (adapted from Ploeg et al 2006)	283
	Appendix 4: Revised quality assessment checklist	284
	Appendix 5: Details of studies which were potentially relevant to the review of clinical effectiveness copies of which could not be obtained within the study timescale	287
	Appendix 6: Details of included studies relating to trials which met the inclusion criteria for the review of clinical effectiveness	289

Appendix 7: Table of excluded studies with rationale 294
Appendix 8: Data abstraction tables 296
Appendix 9: Clinical efficacy data 309
Appendix 10: Registry data 346
Appendix 11: Longitudinal pain trends 351
Appendix 12 Review of observational studies: estimating mortality differences between treatments for vertebral compression fractures 355

APPENDIX 13 MTC of mean difference in VAS during stable period 367
11. REFERENCES 374

LIST OF TABLES & FIGURES

Table 1 Observed and expected survival following vertebral fracture in men and women aged ≥65 years (data from van Staa et al 2001) 14
Table 2 Principal sources for each trial 61
Table 3 Ongoing or unpublished randomised trials of percutaneous vertebroplasty or percutaneous balloon kyphoplasty in patients with painful osteoporotic vertebral compression fractures 63
Table 4 Characteristics of included studies 65
Table 5 Baseline demographic data 68
Table 6 Crossover from allocated treatment groups 78
Table 7 Numbers of potential participants who refused to participate in the included studies 84
Table 8 Mean Dallas Pain Questionnaire scores, by domain, before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: medium- and long-term outcomes: data from Rousing et al 89
Table 9 Use of walking aids: data from the FREE study 96
Table 10 Bed rest due to back pain in the previous 14 days: data from the FREE study 96
Table 11 Number of patients in the Buchbinder and INVEST studies showing improvement in pain scores at one month: data from Staples et al 2011 102
Table 12 Number of patients using opioids before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures 104
Table 13 Number of treated vertebrae with imaging-identified cement leakage 109
Table 14 Number of patients suffering new incident radiographic vertebral fractures 112
Table 15 Incidence of clinical vertebral fractures 114
Table 16 Adverse events reported from the study by Buchbinder et al 115

Table 17	The FREE study: patients with serious adverse events, data to 24 months	116
Table 18	Effect of treatment on pain at 1 month in the INVEST study, by duration of pain at baseline (data from Kallmes et al 2009)	117
Table 19	Change from baseline in mean (SD) RDQ scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	117
Table 20	Change from baseline in mean (SD) EQ-5D scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	118
Table 21	Change from baseline in mean (SD) overall pain scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	118
Table 22	Change from baseline in mean (SD) RDQ scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	119
Table 23	Change from baseline in mean (SD) EQ-5D scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	119
Table 24	Change from baseline in mean (SD) overall pain scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011)	120
Table25	Adverse events in large (n≥200) balloon kyphoplasty case series	125
Table 26	Adverse events in large (n≥200) vertebroplasty case series	128
Table 27	Summary of results estimating a mortality benefit associated with BKP or PVP	132
Table 28	Complications requiring revision surgery following vertebroplasty (data from Yang et al)	137
Table 29	Case reports of pulmonary embolism after vertebral augmentation	138
Table 30	Number of patients with asymptomatic cement leaks	151

Table 31	Number of treated vertebrae with asymptomatic cement leaks	151
Table 32	The VAS scores from Johnson and Johnson's network meta analysis including OPLA	157
Table 33	Odds for cement leakage by treatment	158
Table 34	Comparison of refracture rates by intervention	159
Table 35	Resource use in the preliminary, operating and post-operative phases	161
Table 36	Average procedure length and split of levels	162
Table 37	Estimated costs associated with vertebroplasty	162
Table 38	The assumed length of stay, cost per day and total cost assumed in the Johnson and Johnson submission	163
Table 39	The base case deterministic results in the Johnson and Johnson submission	164
Table 40	The base case deterministic results in the scenario analyses within the Johnson and Johnson submission	165
Table 41	A threshold analysis on key parameters affecting the deterministic results in the Johnson and Johnson submission	166
Table 42	The assumed undiscounted QALYs gained per treatment in the Medtronic submission basecase	169
Table 43	Procedure Costs reported in the Medtronic Submission	170
Table 44	The assumed length of stay, cost per day and total cost assumed in the Medtronic submission	171
Table 45	The deterministic base case results presented in the Medtronic submission	172
Table 46	Sensitivity Analyses presented in the Medtronic submission on the impact of assumed mortality benefit.	173
Table 47	Sensitivity Analyses presented in the Medtronic submission on the impact of assumed length of stay following BKP	174
Table 48	Sensitivity Analyses presented in the Medtronic submission on the impact of assuming no further benefit beyond the time horizon of the trials	175
Table 49	The probabilistic base case results presented in the Medtronic submission	175
Table 50	The assumed annual risks of vertebral fracture following an initial vertebral fracture based on age and T-Score on entry to the model	181
Table 51	The assumed annual risks of hip fracture following an initial vertebral fracture based on age and T-Score on entry to the model	182
Table 52	The hazard ratios within the three scenarios used to explore the effects of mortality associated with BKP, PVP and OPM	186
Table 53	The acquisition cost of each intervention assumed by the assessment group	201
Table 54	The total preliminary, operating and post-operative costs assumed in the	201

	model	
Table 55	The estimated length of stay in days (standard error) assumed in the manufacturers' submissions	202
Table 56	The base case estimated length of stay assumed by the assessment group	204
Table 57	A comparison of the mathematical models structures developed by the Assessment Group, Johnson and Johnson and Medtronic.	205
Table 58	The basecase aggregated costs assumed by the manufacturers and the assessment group A comparison of the mathematical models developed by the Assessment Group, Johnson and Johnson and Medtronic.	206
Table 59	A comparison of the results produced by the Assessment Group model when using Johnson and Johnson basecase data	207
Table 60	A comparison of the results produced by the Assessment Group model when largely using Medtronic basecase data	207
Table 61	The results of the Assessment Group's foundation analysis assuming no mortality benefit for BKP, PVP or OPLA	212
Table 62	The results of the Assessment Group's foundation analysis assuming a relative risk of mortality ([redacted]).	213
Table 63	The deterministic results produced by Assessment Group - Scenario 1	218
Table 64	The probabilistic results produced by Assessment Group - Scenario 1	223
Table 65	Sensitivity analyses conducted on Scenario 1	224
Table 66	The deterministic results produced by Assessment Group - Scenario 2: FREE data	225
Table 67	The probabilistic results produced by Assessment Group - Scenario 2: FREE data	226
Table 68	Sensitivity analyses conducted on Scenario 2: FREE Data	227
Table 69	The deterministic results produced by Assessment Group - Scenario 2: Buchbinder data	228
Table 70	The probabilistic results produced by Assessment Group Scenario - 2: Buchbinder data	229
Table 71	Sensitivity analyses conducted on Scenario 2: Buchbinder Data	230
Table 72	The deterministic results produced by Assessment Group - Scenario 2: INVEST data	231
Table 73	The probabilistic results produced by Assessment Group - Scenario 2: INVEST data	232
Table 74	Sensitivity analyses conducted on Scenario 2: INVEST Data	233
Table 75	The deterministic results produced by Assessment Group - Scenario 3	234
Table 76	The probabilistic results produced by Assessment Group - Scenario 3	235
Table 77	Sensitivity analyses conducted on Scenario 3	236
Table 78	The deterministic results produced by Assessment Group - Scenario 4: FREE	237

data
Table 79 The probabilistic results produced by Assessment Group - Scenario 4: FREE data 238
Table 80 Sensitivity analyses conducted on Scenario 4: FREE Data 239
Table 81 The deterministic results produced by Assessment Group - Scenario 4: Buchbinder data 240
Table 82 The probabilistic results produced by Assessment Group - Scenario 4: Buchbinder data 241
Table 83 Sensitivity analyses conducted on Scenario 4: Buchbinder Data 242
Table 84 The deterministic results produced by Assessment Group - Scenario 4: INVEST data 243
Table 85 The probabilistic results produced by Assessment Group - Scenario 4: INVEST data 244
Table 86 Sensitivity analyses conducted on Scenario 4: INVEST Data 246
Table 87 The deterministic results produced by Assessment Group - Scenario 5 247
Table 88 The probabilistic results produced by Assessment Group - Scenario 5 249
Table 89 Sensitivity analyses conducted on Scenario 5 250
Table 90 The deterministic results produced by Assessment Group Scenario 6: FREE data 251
Table 91 The probabilistic results produced by Assessment Group - Scenario 6: FREE data 252
Table 92 Sensitivity analyses conducted on Scenario 6: FREE Data 253
Table 93 The deterministic results produced by Assessment Group - Scenario 6: Buchbinder data 254
Table 94 The probabilistic results produced by Assessment Group Scenario 6: Buchbinder data 255
Table 95 Sensitivity analyses conducted on Scenario 6: Buchbinder Data 256
Table 96 The deterministic results produced by Assessment Group - Scenario 6: INVEST data 258
Table 97 The probabilistic results produced by Assessment Group - Scenario 6: INVEST data 259
Table 98 Sensitivity analyses conducted on Scenario 6: INVEST Data 260
Table 99 The effect on the cost per QALY gained of high viscosity cement when changing both the QALY gained and the level of re-operations 265

Figure 1 Clinical Effectiveness: Summary of Study Selection and Exclusion 59
Figure 2 Risk of bias summary: review authors' judgements about each risk of bias item for each included study (+ = low risk; - = high risk; ? = unclear risk) 75
Figure 3 External validity and precision summary: review authors' judgements about 86

	each included study (+ = good generalisability/precision; - = poor generalisability/precision; ? = unclear generalisability/precision)	
Figure 4	EQ-5D data recorded in Buchbinder et al	91
Figure 5	EQ-5D data recorded in the FREE trial	91
Figure 6	EQ-5D data recorded in the INVEST trial	92
Figure 7	EQ-5D data recorded in the Rousing trial	92
Figure 8	Longitudinal pain reduction trends in vertebroplasty without AIC data	99
Figure 9	Longitudinal pain reduction trends in vertebroplasty with AIC data	99
Figure 10	Longitudinal pain reduction trends in balloon kyphoplasty	100
Figure 11	Longitudinal pain reduction trends in OPLA excluding AIC data	100
Figure 12	Longitudinal pain reduction trends in OPLA including AIC data	100
Figure 13	Longitudinal pain reduction trends in OPM	101
Figure 14	Overall mortality at 12 months	107
Figure 15	Patients with new incident radiographic vertebral fractures at 12 months	113
Figure 16	The cost-effectiveness plane and the cost-effectiveness frontier associated with the base case deterministic results in the Johnson and Johnson submission	164
Figure 17	A tornado plot of univariate sensitivity comparing PVP with NIM in the base case	166
Figure 18	The diagrammatic representation of the Medtronic model	167
Figure 19	The CEAC presented in the Medtronic submission	176
Figure 20	A scatter plot of the paired BKP and PVP results	176
Figure 21	Diagram of the conceptual model	180
Figure 22	An illustrative example of methodology regarding VAS scores post-intervention	187
Figure 23	The relationship between initial VAS score and stable VAS score	188
Figure 24	The relationship between the mean initial VAS score and the difference in the stable VAS score	189
Figure 25	The relationship between the mean initial VAS score and the difference in the stable VAS score PVP vs OPM trials only	190
Figure 26	The relationship between the difference in the initial VAS score and the difference in the stable VAS score	190
Figure 27	The relationship between the difference in the initial VAS scores and the difference in the stable VAS scores using only the OPM vs PVP trials.	191
Figure 28	The network of evidence regarding VAS scores	191
Figure 29	The relative VAS scores	192

Figure 30	A plot of absolute VAS versus absolute EQ-5D	193
Figure 31	A plot of absolute VAS versus absolute EQ-5D excluding INVEST data	194
Figure 32	A plot of absolute RDQ versus absolute EQ-5D	194
Figure 33	A plot of absolute RDQ versus absolute EQ-5D excluding INVEST data	195
Figure 34	An illustrative example of the conversion of VAS scores to EQ-5D	196
Figure 35	The different assumed hospitalisation costs within the manufacturers submission and Assessment Group base case	203
Figure 36	Derivation of the Assessment Group's six scenarios.	211
Figure 37	Univariate analyses regarding patients' characteristics	214
Figure 38	Univariate analyses regarding hospitalisation costs, operation cost and cement costs.	210
Figure 39	Univariate analyses regarding the costs of equipment required for OPLA and the cost of the procedure when undertaking OPLA	215
Figure 40	Univariate analyses regarding discount rates, bisphosphonate usage and bisphosphonate wane period	216
Figure 41	Univariate analyses regarding the assumed time of convergence, the trials used in the VAS to EQ-5D mapping and the inclusion of adverse events associated with treatment	217
Figure 42	Univariate analyses using change in EQ-5D data directly from trials	219
Figure 43	Univariate analyses regarding mortality and fracture rates	219
Figure 44	A plot of the deterministic results produced by Assessment Group -Scenario 1	222
Figure 45	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 1	223
Figure 46	A plot of the deterministic results produced by Assessment Group - Scenario 2: FREE data	225
Figure 47	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 FREE data	226
Figure 48	A plot of the deterministic results produced by Assessment Group - Scenario 2: Buchbinder data convergence starts at 24 months	228
Figure 49	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 Buchbinder data	229
Figure 50	A plot of the deterministic results produced by Assessment Group - Scenario 2: INVEST data convergence starts at 24 months	231
Figure 51	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 INVEST data	232
Figure 52	A plot of the deterministic results produced by Assessment Group - Scenario 3	234
Figure 53	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 3	235

Figure 54	A plot of the deterministic results produced by Assessment Group - Scenario 4: FREE data	237
Figure 55	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 4 FREE data	238
Figure 56	A plot of the deterministic results produced by Assessment Group - Scenario 4: Buchbinder data convergence starts at 24 months	240
Figure 57	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 4 Buchbinder data	241
Figure 58	A plot of the deterministic results produced by Assessment Group - Scenario 4: INVEST data convergence starts at 24 months	244
Figure 59	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 4 INVEST data	245
Figure 60	A plot of the deterministic results produced by Assessment Group - Scenario 5	248
Figure 61	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 5	249
Figure 62	A plot of the deterministic results produced by Assessment Group Scenario 6: FREE data	251
Figure 63	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 FREE data	252
Figure 64	A plot of the deterministic results produced by Assessment Group - Scenario 6: Buchbinder data convergence starts at 24 months	254
Figure 65	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 Buchbinder data	255
Figure 66	A plot of the deterministic results produced by Assessment Group - Scenario 6: INVEST data convergence starts at 24 months	258
Figure 67	The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 INVEST data	259
Figure 68	An exploratory analysis of effect of assuming additional QALY gains associated with high-viscosity cement compared with low-viscosity cement.	263
Figure 69	An exploratory analysis of the cost per QALY gained of PVP versus OPM assuming that the response of OPLA could be generated in OPM patients through education. An exploratory analysis of effect of assuming additional re-operations associated with high-viscosity cement compared with low-viscosity cement.	264
Figure 70	An exploratory analysis of effect of assuming additional re-operations associated with high-viscosity cement compared with low-viscosity cement.	267

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader.

1. DEFINITION OF TERMS

Bone mineral density	A measure of the strength of the bones, ascertained by calcium content
Kyphosis	Abnormal curvature of the spine
Minimal clinically important difference	The smallest change in an outcome measure which reflects a change in symptom which can be considered important to patients
Osteopaenia	A condition in which bone density is lower than the average in the healthy young population. Diagnosis requires a T-score between -1.0 and -2.5.
Osteoporosis	A severe loss of bone mineral density and deterioration of bone microarchitecture. Diagnosis requires a t-score below -2.5.
Parapesia	Motor weakness, especially of the legs.
Radiculopathy	Pressure on, or other damage to, a nerve root.
T-score	The number of standard deviations of an individual's bone density above or below the bone mineral density of a healthy 30-year-old matched for gender and ethnicity.
Vertebral augmentation	The addition of cement into a vertebra affected by a compression fracture in an attempt to stabilise it, and in some cases also to reduce the compression. Vertebral augmentation is a generic term which embraces both percutaneous vertebroplasty and percutaneous balloon kyphoplasty.
Z-Score	The number of standard deviations that a woman is from the average bone mineral density of women of the same age.

LIST OF ABBREVIATIONS

AE	Adverse event
AG	Assessment Group
AQoL	Assessment of Quality of Life scale
BKP	Balloon kyphoplasty
BMD	Bone mineral density
CEAC	Cost-effectiveness acceptability curve
CHC	Carbonated hydroxyapatite cement
DPQ	Dallas Pain Questionnaire
EQ-5D	EuroQol- 5 dimensions scale
HES	Hospital Episode Statistics
HR	Hazard ratio
HRT	Hormone replacement therapy
IPD	Individual patient data
MCID	Minimal clinically important difference
MMSE	Mini Mental State Examination
NMB	Net Monetary Benefit
NSAID	Non-steroidal anti-inflammatory drug
ODI	Oswestry Disability Index
OPLA	Operative placebo with local anaesthesia
OPM	Optimal pain management
OR	Odds ratio
QALY	Quality adjusted life year
QUALEFFO	Quality of life questionnaire of the European Foundation for Osteoporosis
PMMA	Polymethylmethacrylate
PVP	Percutaneous vertebroplasty
RCT	Randomised controlled trial
RDQ	Roland-Morris Disability Questionnaire
RR	Risk ratio
SF-36	Medical Outcome Study Short Form 36
SOF-ADL	Study of Osteoporotic Fractures and Activities of Daily Living questionnaire
VAS	Visual Analogue Scale
VBH	Vertebral body height
VCF	Vertebral compression fracture

2. EXECUTIVE SUMMARY

2.1 Background

Osteoporosis is a systemic skeletal disease characterised by low bone mass and microarchitectural deterioration of bone tissue, with a resulting increase in bone fragility and susceptibility to fracture. The clinical significance of osteoporosis lies not in low bone mass per se but in the fractures that may occur as a consequence. In vertebral fracture, one or more vertebrae are compressed, leading to a reduction in height and potentially also to abnormal curvature of the spine (kyphosis). Vertebral compression fractures (VCFs) can lead to severe acute and chronic pain, impaired mobility, and reduced quality of life. They have also been linked to poor cardiopulmonary function and appetite, and an increased risk of mortality. Although VCFs are thought to be common, it is difficult to give a precise estimate of prevalence and incidence, since the majority remain undiagnosed. When painful VCFs do come to clinical attention, they are typically treated with optimal pain management (OPM) consisting of analgesics, bed rest, and back bracing. However, this approach is unsatisfactory for a proportion of patients, and when used as a longer-term treatment, can lead to exacerbation of the underlying osteoporosis.

Percutaneous vertebroplasty (PVP) is a minimally invasive surgical procedure in which bone cement (such as PMMA, glass polymers, hydroxyappetite and calcium phosphate) is injected into a fractured vertebra under radiological guidance using fluoroscopy. The procedure is usually performed under intravenous sedation or light general anaesthesia. A disposable bone biopsy needle or trocar needle is placed centrally in the vertebral body using an image guided safe access route. This may be done bilaterally through the pedicles, oblique across one pedicle or lateral oblique through the base of the pedicle. The cement is then injected very slowly, again under constant fluoroscopic guidance. Percutaneous balloon kyphoplasty (BKP) is a variation of this approach, in which an inflatable balloon tamp is placed in the collapsed vertebra prior to cement injection in order to create a cavity allowing low pressure injection. A potential advantage of kyphoplasty is that it may partially correct the reduction in vertebral height however the degree of height restoration may be none or minimal. Early case reports, retrospective case series and quasi-experimental studies suggested these procedures led to dramatic improvements in pain and physical functioning. Furthermore, there are plausible biomechanical reasons that may account for these improvements, such as stabilisation of the collapsed vertebra, correction of kyphotic deformity, and height restoration. However, two recent double-blind, placebo controlled trials of PVP suggest that the procedure may provide no greater benefits than administration of local anaesthetic to the affected area.

2.2 Objectives

The objective of this review was to systematically evaluate and appraise the clinical and cost-effectiveness of percutaneous vertebroplasty and percutaneous balloon kyphoplasty in reducing pain and disability in people with osteoporotic vertebral compression fractures in England and Wales. The study also included a narrative review of safety.

2.3 Methods

A systematic search of databases including Medline; CINAHL; EMBASE; EconLit; the Cochrane Library, and DARE was conducted. Search terms included 'vertebroplasty', 'kyphoplasty', and a broad variety of related clinical terms. Studies met the inclusion criteria if they were randomised controlled trials (RCTs) including people of any age and either gender with painful osteoporotic vertebral compression fractures. The intervention groups of these trials must have received PVP or BKP, and the comparators were the interventions themselves, conservative management, or defined as sham surgery. Primary outcomes were: health-related quality of life, back-specific functional status/mobility, pain/analgesic use, vertebral body height and angular deformity, incidence of new vertebral fractures, and progression of treated fracture. Safety was assessed in a narrative review including data from the RCTs of PVP and BKP, along with large case series ($\geq$200) and individual case reports of complications.

Data were extracted independently by two reviewers using a standardised data extraction form; discrepancies were resolved by discussion. The quality of the included studies was critically assessed by the same two reviewers using a tool based on the criteria proposed by the Centre for Reviews and Dissemination and the Cochrane Collaboration for assessing the risk of bias in randomised trials, and also including some vertebral augmentation-specific items.

Due to the potential impact of baseline imbalances in the degree of pain and disability reported by patients with osteoporotic VCFs, outcomes which were reported as continuous data were assessed in terms of the difference between the mean changes from baseline in the intervention and control groups, rather than absolute differences at any time-point. For dichotomous outcomes, relative risks, with confidence intervals and p values, were calculated using the Cochrane Collaboration Review Manager© Software (version 5.1) if such data were not reported by the study investigators. Where appropriate, a meta-analysis was carried out with random effects models, using Review Manager© Software (version 5.1). However, such meta-analysis was limited to dichotomous outcomes. It was not considered appropriate to

undertake a meta-analysis of continuous or quasi-continuous outcomes because a previous meta-analysis of individual patient data from the two double-blind placebo controlled trials has already been published. Where meta-analysis was not possible, published data were tabulated and discussed in a narrative review.

Medtronic provided observational data indicating that vertebral augmentation may be associated with a beneficial mortality effect, and that potentially BKP was more efficacious than PVP. The clinical hypothesis for this effect is that as patients become more mobile more quickly, the typically elderly patients are less prone to infection. These data were formally critiqued.

A mathematical model was constructed to explore the cost-effectiveness of BKP, PVP (using low-viscosity cement in 85% of patients and high-viscosity in 15% of patients) and Operative placebo with local anaesthesia (OPLA) compared with OPM. Due to uncertainty in the evidence base, six scenario analyses were conducted that assessed combinations of assumptions on mortality (differential beneficial effects for BKP and PVP; equal beneficial effects for BKP and PVP; and no effect assumed) and derivation of utility data (either solely mapped from Visual Analogue Scale (VAS) pain score data produced by a network meta-analysis or using direct EQ-5D data from the trials). Extensive sensitivity analyses were conducted on each of the six scenarios. Exploratory analyses were conducted on the cost-effectiveness of using high-viscosity cement in all patients, on the available costs for patient education to obtain the OPLA response whilst maintaining a cost per QALY gained ratio below £20,000 and on the use of initial facet joint injections.

2.4 Results

- Number and quality of studies

28 articles relating to a total of nine RCTs were identified and included in the review of clinical effectiveness. This body of literature was of variable quality, with the two double-blind, placebo-controlled trials (Buchbinder et al, INVEST) being at the least risk of bias. The most significant methodological issue among the remaining trials was lack of blinding for both study participants and outcome assessors. In addition, only the two placebo-controlled trials provided adequate information on the prior training, skills, and knowledge of the operators.

- Summary of benefits and risks

Broadly speaking, the literature suggests that both PVP and BKP provide substantially greater benefits than OPM in open label trials. However, in double-blinded trials PVP was shown to

have no more benefit than local anaesthetic; no trials of BKP compared with local anaesthesia have been conducted.

Quality of life was most often assessed with the EuroQol- 5 dimensions (EQ-5D) and / or Quality of life questionnaire of the European Foundation for Osteoporosis (QUALEFFO) scales. Findings indicated greater improvements on both these measures in the open label trials of PVP (Blasco, Rousing, Farrokhi, VERTOS, VERTOS II); however, no differences in quality of life were observed in either of the placebo-controlled, double-blind trials (Buchbinder, INVEST). Four open-label studies (Farrokhi, FREE, Rousing, VERTOS II), found significantly greater improvements in pain among the operated cohorts, while the double-blind trials found no or a small non-significant benefit. Although there was a trend toward greater pain reduction in the PVP group in one of these placebo-controlled trials (the INVEST study), this may have been confounded by a higher level of opioid use among the PVP group. With respect to analgesic use too, there were greater reductions among nonoperated patients in the open-label trials, while no significant between-group differences were seen in the double-blind trials. In a head to head trial of PVP and BKP (Liu) VAS pain scores did not differ significantly between the treatment groups.

There were no data on restoration of vertebral body height or kyphotic wedge angle that could be compared between studies. However, the one trial that undertook a comparison of PVP and BKP (Liu) suggests that BKP may be the more effective method. Only one study comparing BKP with OPM was identified (FREE). This suggested that BKP is more effective for reducing pain, and improving back-related functional ability and quality of life. However, the methodological limitations of this study – most notably lack of blinding and unexpected imbalances in dropout – made it difficult to draw inferences with any confidence.

Known complications of PVP and BKP include pulmonary embolism, periprocedural hypotension, radiculopathy, damage to surrounding tissue, paraparesia, paraplegia, rib fracture, and postoperative infection. Most of these complications are associated with leakage of bone cement outside the treated vertebra. Although intradiscal leakage is unlikely to lead to complications, epidural leakage can have serious consequences, and a number of procedure-related deaths have been reported. Incidence of serious complications is rare, but the long-term implications of clinically silent cement leakages and pulmonary emboli remain poorly understood.

A meta-analysis of mortality rates suggested that PVP might be associated with reductions in mortality. However, this effect failed to reach statistical significance and the included trials

were not designed to detect this outcome. A formal analysis of mortality data undertaken within this report concludes that it is possible that there is a causal difference in mortality between patients treated using OPM and patients receiving BKP or PVP given the size of the effect. Appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. It is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment based on the data presented. There is also considerable uncertainty were BKP and PVP assumed to have a mortality benefit, in whether OPLA would also produce a mortality benefit.

The cost-effectiveness ratios of the interventions were driven by the scenario chosen. If a differential effect was chosen, then BKP consistently had a cost per QALY gained ratio below £20,000. If a pooled beneficial effect was used then PVP consistently had a cost per QALY gained ratio below £10,000. Where no mortality effect was assumed then the derivation of utility influenced the results. Using solely the EQ-5D values mapped from VAS pain scores produced by a network meta-analysis PVP typically had a cost per QALY gained ratio below £20,000 with the exception of when a number of parameters were altered that did not favour PVP. However, when data from the two high-quality blinded trials (Buchbinder et al or INVEST) were assumed appropriate then the cost per QALY gained ratios for PVP and BKP were often greater than £20,000 dependent on other assumptions made. The exploratory analyses indicated that: the use of high-viscosity cement in all patients was unlikely to have a cost per QALY below £20,000; that potentially sums in excess of £500 (and potentially considerable more) per patient could be spent to achieve the OPLA response rather than undertake PVP and that an initial facet joint injection prior to vertebral augmentation appeared a sensible option.

2.5 Discussion

- Strengths, limitations of the analyses and uncertainties

To our knowledge, this is the first systematic review to undertake a comprehensive clinical and cost effectiveness analysis of PVP and BKP for the treatment of osteoporotic VCFs. The clinical effectiveness analysis included RCTs only, and provided an overview of the complications that may arise from these procedures. However, the internal validity of the included literature was compromised by widespread lack of blinding. To date, there has only been one open-label trial to compare BKP with conservative management, so the effectiveness of this procedure was particularly difficult to establish. The use of subjective ratings of pain as an outcome measure may be confounded by various psychosocial and patient-level factors. Important questions that are yet to be convincingly addressed include the

effect of vertebral augmentation on mortality, and on correction of vertebral body height and kyphotic deformity. The analyses conducted the most robust mapping of VAS to EQ-5D of which we are aware, and undertook a network meta-analysis of the VAS data. Extensive scenario and sensitivity analyses were conducted to explore a wide range of different assumptions. Insufficient evidence, particularly on the impact of BKP, PVP and OPLA on mortality rates, means that no definitive conclusion can be made.

- Generalisability of the findings

This review was specific to the population of people with painful osteoporotic VCFs. Hence, the results are not necessarily generalisable to VCFs of other origins (e.g. multiple myeloma, traumatic, metastatic deposits). Most studies did not present data on the ethnic composition of their samples, nor discussed the implications of this for generalisability. Furthermore, the procedures reported in those studies were usually performed by experienced personnel, and therefore their results may differ from those obtained by less experienced practitioners. On the other hand, the age and gender makeup of the study samples was fairly representative of the wider population of people with osteoporotic VCFs. A higher proportion of females took part in the trials (typically around 70%); and the mean sample age was usually early to mid 70s.

2.6 Conclusions

For people with painful osteoporotic VCFs refractory to analgesic treatment, PVP and BKP perform significantly better in unblinded trials than OPM in terms of improving quality of life and reducing pain and disability. However, there is as yet no convincing evidence that either procedure performs better than OPLA with data from two high-quality trials (Buchbinder, INVEST). It can be argued that these procedures should not be undertaken unless the patient has failed to respond to a facet joint injection.

It is possible that BKP and PVP may lead to longer-term reductions in mortality and at different levels of effect; however, this possibility was derived from registry data and without information on the causes of death in these cohorts, and in the absence of randomisation, it was not possible to conclusively establish a causal link. There were no data to analyse whether OPLA would also be associated with mortality benefits. If such benefits exist then the cost per QALY gained of the interventions compared with OPM would be low.

Although complications associated with PVP and BKP are rare, they can be serious, and procedure-related deaths have been reported.

- Suggested research priorities

- There is yet to be a double-blind, placebo controlled trial of BKP. A well designed study comparing BKP with OPLA should be considered.
- There are questions as to whether postoperative pain and quality of life improvements from PVP and BKP arise from a placebo response or the specific efficacy of the procedures. It may be that the failure of PVP to demonstrate greater benefits than OPLA suggests placebo efficacy only. Alternatively, it may be that the infusion of local anaesthetic has specific mechanisms of efficacy over conservative treatment. RCTs comparing local anaesthesia with OPM, and multi-arm RCTs comparing vertebral augmentation, local anaesthesia, facet joint injection, patient education and OPM would provide useful data.
- The effect of vertebral augmentation on mortality is an important, yet inadequately understood issue. Large-scale registry data from Germany and the USA suggests that people with osteoporotic VCFs who have received augmentation have significantly improved survival rates however a definitive causal link could not be established. The effect of augmentation on mortality, and the impact of various extraneous variables, should be investigated through further retrospective case series with more details on causes of death. Ideally, this outcome would be explored in a well controlled RCT. However, the sample size and length of follow-up required to detect meaningful differences would make such a trial difficult to perform.
- The length of stay associated with patients receiving OPM, PVP and BKP is not known with certainty, with the pivotal trials suggesting that the length of stay is considerably shorter than hospital database values. A prospective study to record such values would be beneficial.
- Saggital balance and spinal deformity have a substantial impact on quality of life and fracture-related disability. However, the effectiveness of PVP and BKP in restoring these morphometric parameters is yet to be studied in high quality studies.

3. BACKGROUND

3.1. Description of health problem

Aetiology

Osteoporosis is a systemic skeletal disease characterised by low bone mass and microarchitectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture.[1] A definition of osteoporosis has been developed based on bone mineral density (BMD), as this can be measured with precision and accuracy. This defines osteoporosis in terms of the T-score, the number of standard deviations (SD) by which the individual's BMD, as measured by dual x-ray absorptiometry (DXA) at the lumbar spine, hip (total hip or femoral neck) and forearm, differs from the average BMD of healthy young women. The BMD osteoporosis threshold proposed for Caucasian women is a T-score of 2.5 SD or more below that average (i.e. a T-score $\leq$-2.5); a T-score of between 1 and 2.5 SD below that average (i.e. -1 to -2.5) indicates osteopenia.[2]

The clinical significance of osteoporosis lies not in low BMD per se but in the fractures that may occur as a consequence of low BMD: without a fracture, a person suffering from osteoporosis will not suffer morbidity. Fractures are considered to be osteoporotic if they occur in a person with low BMD as a result of little or no trauma – the equivalent of a fall from standing height or less.[3] Vertebral fractures are amongst the most common osteoporotic fractures. The risk of such fractures approximately doubles with each SD decrease in lumbar spine BMD.[4] However, as the occurrence of a VCF, even if asymptomatic, increases the risk of further VCFs by at least four-fold independently of BMD, there appears to be another aspect of bone fragility which is not measured by bone densitometry.[4] Research in women with postmenopausal osteoporosis indicates that, in the absence of antiosteoporotic medication such as bisphosphonates, once a VCF has occurred, the risk of a subsequent VCF occurring within a year is about 19% (95% CI 13.6 to 24.8%).[5] Although about a quarter of VCFs result from falls, most are associated with routine daily activities such as bending or lifting light objects.[6]

In vertebral fracture, or vertebral compression fracture (these terms are used interchangeably within the literature), the vertebra is compressed, leading to a reduction in its height and potentially also to abnormal curvature of the spine (kyphosis). However, there is no universally accepted definition of a vertebral compression fracture (VCF). Definitions which depend on a reduction in the height of an individual vertebral body, whether relative or absolute, are restricted in their utility by the need for an earlier image against which to

identify the change; they are therefore most commonly used in research studies. The same is true of the most widely accepted definition of VCF, Genant's semiquantitative method, which classifies changes in vertebral body shape in terms of reductions in overall height and area.[7] In the absence of an earlier image, the reduction in vertebral height may be assessed by comparison with an adjacent undeformed vertebra. VCFs which are only identified on x-rays taken for research, population screening, or other purposes are termed radiographic or morphometric fractures.

Some osteoporotic VCFs are diagnosed clinically, usually when a person presents with back pain and a subsequent x-ray is interpreted as showing a fracture to a vertebral body. However, accurate clinical diagnosis of a new VCF may be confounded by the high prevalence of back pain from other causes, by changes in vertebral morphology which are either longstanding or due to causes other than fracture, or by nonstandardised interpretation of spinal x-rays.[8] The evidence from clinical trials in which VCFs are identified radiographically suggests that about two-thirds of VCFs are not brought to clinical attention.[9] This may be because the fractures are associated with no, or only mild, symptoms, or because any symptoms are attributed to another cause, such as muscle strain.[10] Previously unreported fractures may be identified only when they have caused kyphosis and obvious loss of height.[11] However, kyphosis also occurs in osteoporotic women without VCFs, and in these women it is presumably due to non-skeletal factors such as poor muscle tone and loss of disc height by degenerative spondylosis.[12]

Research has shown that women with previously unreported vertebral deformities which are found incidentally during population screening are substantially more likely to have chronic back pain and functional difficulties than women without vertebral deformities. However, women with clinically diagnosed fractures are more likely to report symptoms than those whose fractures are only detected by population screening.[10] Only those patients who present to healthcare professionals with clinical VCFs and severe pain are likely to be considered for percutaneous vertebroplasty (PVP) or balloon kyphoplasty (BKP). Cummings and Melton[4] have suggested that fewer than 10% of radiographically-detected fractures – i.e. at most a third of clinical fractures – are severe enough to require hospital admission. However, patients with fractures of such symptomatic severity are presumably those who are most likely to be offered PVP or BKP.

Osteoporotic VFCs may be due to primary or secondary osteoporosis. Primary osteoporosis is defined as osteoporosis which is not associated with any other illness; it is generally associated with aging, and is particularly common in postmenopausal women. Secondary

osteoporosis may be related to certain medical conditions (e.g. hyperthyroidism, malabsorption and extreme dieting) or to prolonged steroid therapy.[13] Most osteoporotic VCFs occur in women with primary postmenopausal osteoporosis. However, because of the increase in chronic steroid use, the incidence of VCFs due to secondary osteoporosis is increasing.[14]

The short-term impact of VCFs

Clinical VCFs can cause considerable acute pain, which may be persistent. This pain is exacerbated by movement and reduced by rest, and may therefore limit mobility;[14,15] consequently, particularly severe cases may require hospitalisation.[16] Radiculopathy (pressure on, or other damage to, the nerve root) is not uncommon, and may cause either unilateral or bilateral pain radiating along the affected nerve.[15] Such acute pain is intense at the fracture site; it usually lasts 4 to 6 weeks. This is illustrated by data from the VERTOS II study: the study inclusion criteria specified that participants should have had back pain for no more than 6 weeks, and 53% (229/431) of people who initially appeared to be eligible for randomisation became ineligible during the course of the screening process (i.e. in less than 6 weeks from pain onset) because of spontaneous pain relief.[17]

However, in some patients the acute pain associated with a VCF is followed by chronic pain. This often occurs either when one vertebra is particularly severely compressed, or when multiple vertebrae are fractured.[18] It may be predominantly caused not by the fracture itself but by strain on muscles and ligaments secondary to kyphosis, and therefore tends not to respond to the management strategies used for acute pain (rest, activity modification, and local and/or systemic analgesics) but may be better addressed through exercise.[10]

Investigators have sought means of differentiating patients in whom pain following VCF is likely to resolve relatively quickly from those who are likely to develop chronic pain. Klazen et al studied conservatively-treated patients with a radiographically-diagnosed VCF who had had pain for no more than 2 weeks. By 6 months the mean pain score had decreased significantly (i.e. by 50% or more) from baseline, but no significant decrease was seen between 6 and 23 months; thus 63% of patients (22/35) had significant pain relief at 6 months, but the proportion had only increased to 69% (25/36) at 23 months. The patients could be divided into two categories: in those with significant pain relief at 23 months, a rapid decline in pain in the first 6 months continued more slowly thereafter, whereas in those without significant pain relief at 23 months, after a small decrease in pain in the first 6 months, there was no further decrease in pain, which might even increase. None of the recorded baseline factors (age, gender, number of VCFs at baseline, conservative therapy

frequencies, grade of VCF, or pain medication) predicted significant pain relief at 6 or 23 months, but a high pain score at 6 months predicted no significant pain relief at 23 months (OR 0.254, 95% CI 0.293 to 0.938, p=0.030).[19] However, in a study of osteoporotic postmenopausal women with acute back pain, Lyritis et al found that those with radiological evidence of a fully collapsed vertebra which was considered responsible for the pain had pain which was severe (9±0.2 on a scale of 0-10) but of short duration (4-8 weeks). By contrast, women with radiological evidence of only a mild fracture, or with no radiological signs of fracture, had on average three attacks of pain, representing gradual fracture progression; thus, the intensity of the pain was less (6±1.8), and the initial attack was of shorter duration, but the time to final resolution was longer (6-20 months).[11]

The picture is complicated by the fact that it can be difficult to determine the precise date of occurrence of a vertebral compression fracture. In some cases, following the sudden onset of back pain, conventional radiographs cannot identify a vertebral deformity but scintigraphic imaging may identify a "hot spot" which appears as a typical compression fracture on subsequent radiographs; in other cases, the patient may be identified as having an acute vertebral fracture when the deformity may in fact be seen on earlier radiographs. Moreover, the occurrence of additional episodes of pain associated either with new fractures or with the progression of the original deformity may make it difficult to determine the duration of pain associated with a specific fracture.[10]

The longer-term impact of VCFs

Patients who have suffered one VCF are not only at risk of developing chronic pain but also at increased risk of suffering another VCF. They are thus also at risk of long-term morbidity caused by the back pain and progressive loss of height and kyphosis associated with multiple fractures, and this in turn may lead to a loss of mobility which will exacerbate the underlying osteoporosis and increase the risk of future fractures.[6]

People who have suffered a VCF have higher mortality rates than people of the same age who do not have VCFs. Van Staa et al used data from the General Practice Research Database to compare observed and expected survival in England and Wales in men and women aged 65 and over following vertebral fracture.[20] As these fractures had been recorded in the patients' medical records, presumably most if not all were clinical rather than radiographic; given the age group being studied, it seems likely that the majority were osteoporotic. A statistically significant excess of mortality was seen in both genders for up to 5 years following a fracture, but the effect appeared more marked in men than in women (see Table 1).

Table 1: Observed and expected survival following vertebral fracture in men and women aged ≥65 years (data from van Staa et al 2001[20])

	Men		Women	
	Observed	**Expected**	**Observed**	**Expected**
At 3 months	87.8%	97.9%	94.3%	98.4%
At 12 months	74.3%	91.8%	86.5%	93.6%
At 5 years	42.1%	64.4%	56.5%	69.6%

It has been suggested that the primary reason for the excess mortality associated with VCFs is the impact on lung function;[16] abdominal dysfunction associated with kyphosis may also be a contributory factor.[21] Research has shown that pulmonary function is significantly reduced in patients with primary osteoporosis and vertebral fracture, but not in patients with chronic low back pain without evidence of manifest spinal osteoporosis[22] or in healthy controls of the same age.[23] A significant association has been found between the number of vertebral fractures and decline in lung function.[24] However, the increased risk of death may also be due, at least in part, to the co-existence of serious underlying diseases in many individuals with VCF.[25] Thus, research carried out in Sweden found that hospitalisation for vertebral fracture (including traumatic fracture) in men and women aged 50 or over was associated with an increase in the relative risk of death compared with the age- and sex-matched population. However, as the risk was particularly high in the younger individuals and decreased with age, it was suggested that this phenomenon might be related to the impact of trauma injuries or other significant comorbidity and secondary causes of osteoporosis.[26] It is also possible that the complications associated with long-term opioid analgesic use such as respiratory depression, anorexia, and bowel obstruction associated with constipation, contribute to excess mortality.

Similarly, in the USA, a retrospective study was carried out in all residents of Rochester, Minnesota, who had been diagnosed with one or more clinical vertebral fractures between 1985 and 1989; the maximum follow-up appears to have been 5 years, and the mean around 2.4 years. This study found that survival was significantly impaired in the short- to medium-term in the 276 patients who experienced fracture following mild-to-moderate trauma (defined as less than, or equal to, a fall from standing height), and whose fractures were not associated with primary or metastatic cancer or localised bone disease. The most commonly reported causes of death in such patients were cardiovascular diseases (43% - mainly coronary artery disease) and malignancies (18%); the mortality due to coronary artery disease or stroke was not higher than expected, but mortality due to cancer and other causes was

elevated. However, relative survival data were presented for all people with clinical vertebral fractures (i.e. including fractures associated with severe trauma or in areas of bone affected by primary or metastatic cancer or localised bone disease). Cooper et al note that the gradual divergence of observed from expected survival suggests that the impaired survival is unlikely to result from the vertebral fracture per se, but is more likely to be due to an indirect association with comorbid conditions which lead to an increased risk of death, with the fractures simply representing a marker of increased frailty.[27]

In 1991, Browner et al published data from the Study of Osteoporotic Fractures (SOF) indicating that, in elderly women, osteopaenia was associated with an elevated risk of non-trauma mortality, especially deaths from stroke.[28] Subsequently, analysis of data from the FIT trial found that, in primarily healthy postmenopausal Caucasian women with osteoporosis or osteopenia, the age-related relative risk of dying following a clinical vertebral fracture was 8.64 (95% CI 4.45-16.74). Despite the fact that only 122 women died during the follow-up period of 3 to 4 years (99 before suffering a fracture at any site, and only 11 following a vertebral fracture), the risk was clearly elevated (although the confidence intervals were wide), and remained virtually unchanged when adjusted individually for other factors (hypertension, smoking, physical activity, health status, CVD, diabetes, and hip BMD). All 11 deaths following a clinical vertebral fracture occurred within a year of that fracture. However, the authors note that the elevated risk of death following clinical vertebral fracture may reflect an ascertainment bias, whereby women with more medical conditions and poorer health are more likely to receive a diagnosis of clinical vertebral fracture because they would be under greater medical surveillance. They also note that they were unable to estimate whether a death following a fracture was due to the fracture itself or to an underlying medical condition. Consequently, clinical vertebral fractures may be a marker for increased mortality rather than being independently linked to an increased risk of death.[29]

- Incidence and/or prevalence

As Cummings and Melton have noted, it is difficult to establish the total incidence and prevalence of VCFs both because of the lack of a universally accepted definition of VCF and because a substantial proportion of VCFs do not come to clinical attention.[4] However, although structural deformity associated with VCFs might lead to serious morbidity and mortality, it is currently only symptomatic fractures which come to clinical attention that are candidates for PVP or BKP in the UK, and thus only the incidence of clinically-diagnosed fractures is relevant to the current technology assessment. Therefore we have not evaluated the possibility of the early use of BKP to address sagittal balance.

The prevalence of VCFs varies from country to country, and a number of factors – including environment, genetics, availability of diagnostic tests and willingness of radiologists to report fractures – are likely to play a part.[6] It is therefore important, for the current technology assessment, to identify the incidence of clinically-diagnosed VCFs specific to England and Wales. However, as Ström et al note, data on the incidence of clinical vertebral fractures are not available for the UK.[30] Holroyd et al[6] have recently estimated that there are 2,188 hospital admissions a year in England and Wales for vertebral fractures in patients aged 45 and over. While it is not fully clear what data were used to inform this estimate, the most likely source appears to be the UK General Practice Research Database (GPRD), which Ström et al have suggested is likely to incorporate substantial under-reporting of clinical VCFs.[31] Hospital Episode Statistics (HES) data, which relate to hospital admissions and outpatient attendances, appear to provide the most reliable data relating to clinical fractures of sufficient severity to be considered for vertebral augmentation. However, Synthes and Medtronic have produced incompatible estimates based on HES data:

- Synthes have estimated, on the basis of 2010/11 HES data, that 20,908 patients a year are diagnosed with osteoporotic VCF in the UK.[32] As the most recent available statistics[33] indicate that the population of England and Wales is approximately 89% of that of the UK as a whole, Synthes' estimate suggests that approximately 18,600 patients in England and Wales are diagnosed with osteoporotic VCF each year; presumably only a proportion of these will then be hospitalised as a result of osteoporotic VCF.
- Medtronic reported HES data indicating that, in 2008/09, 2009/10, and 2010/11, approximately 24,000 patients a year in England and Wales were hospitalised for osteoporotic VCF, while in 2010/11 the total number of patients admitted to hospital for osteoporotic VCFs, vertebral fatigue, or collapsed fractures was 27,051.[34]

Thus, Medtronic's estimate appears to be substantially higher than that of Synthes.

In their sponsor submission, Johnson & Johnson estimated the number of patients per annum in England and Wales who were hospitalised with debilitating pain from osteoporotic VCFs using data from Dr Foster Intelligence, which routinely collects and analyses data from NHS hospitals in England. On this basis, 7073 patients a year were identified as potential candidates for vertebral augmentation.[35] This figure appears to apply to England alone, but this is not wholly clear. It is substantially lower than the figures put forward by Synthes and Medtronic; the submission indicates that this is due to the exclusion of patients with diagnoses other than osteoporosis (e.g. malignancy or trauma),[35] thus making it more relevant to the decision problem.

The ScHARR model uses data on the incidence of vertebral fractures drawn from a different source, a large scale prospective Scottish study.[36] The figures from this study were the basis for a clinical and cost-effectiveness model which has been used in previous NICE assessments of osteoporosis interventions. These are UK specific data and explicitly report vertebral fracture rates rather than relying on estimating these from hip fracture incidence data.

Impact of health problem

- Significance for patients in terms of ill-health (burden of disease).

The significance for patients of VCF falls into three main categories:

- Pain
- Physical changes and impairment
- Psychosocial decline.[10]

These will be discussed in turn below. However, it should be noted that the categories are not entirely independent: pain contributes to physical impairment, and both pain and physical impairment contribute to psychosocial decline.[10]

Pain

VCFs are associated with both acute and chronic pain. The acute pain typically lasts for several weeks or months until the fracture heals. It varies widely in severity, and at worst is described as intolerable; however, it may respond to analgesics. By contrast, chronic pain, which can develop when kyphosis causes strain on muscles and ligaments, often does not respond to analgesics, but may respond to exercises which increase the tone and strength of the back muscles.[10,37]

In a small case-control study, Lyles et al found that pain, as measured by the West Haven-Yale Pain Inventory, was significant worse in women with VCFs than in matched controls (p=0.001).[38]

Physical and functional outcomes

VCFs, and in particular multiple fractures, are associated with decreases in stature and progressive kyphosis which cause loss of lung volume and loss of appetite.[10,24] In the USA, a prospective cohort study of women aged 65 or over found that severe kyphosis was related to pulmonary deaths (HR 2.6, 95% CI 1.3-5.1).[39] However, Ettinger et al found that, in a sample of 610 white women aged 65 to 91, despite greater spinal curvature and height loss, the 10% with the most severe thoracic kyphosis did not report significantly greater back pain or back-related disability, or consider themselves to have poorer health, than the other women.[40]

Vertebral fracture can also lead to a loss of spinal mobility, which causes problems with the activities of daily living. If the fracture is accompanied by acute pain which limits physical activity, this may lead to muscle weakness which may in turn contribute to chronic pain.[10] The rate of decline in BMD also appears to decrease with physical inactivity, and may decrease by as much as 40% during bed rest or post-fracture recovery, thus greatly increasing the risk of subsequent fractures.[10]

The preservation of independence in elderly community-living individuals depends substantially on the extent to which they are able to perform everyday activities such as shopping and preparing meals.[41] A number of studies have found an association between symptomatic VCF and problems with such activities. In small studies, Cook et al found that over 80% of postmenopausal women with a diagnosis of chronic back pain due to osteoporotic VCF reported problems with physical functioning and activities of daily living,[42] while Lyles et al found that women with two or more confirmed VCFs were significantly more likely than age- and race-matched controls with equivalent comorbid conditions to report pain and difficulty in performing functional activities, and to say that their health problems interfered with their daily activities ($p=0.002$).[38] Moreover, a population survey of 1010 white community-dwelling Californian women aged 55 and over found that those with clinically-diagnosed osteoporotic VCFs were significantly more likely to report difficulty in activities such as lifting, shopping, and cooking meals than women without known vertebral fractures (adjusted OR 3.42 (95% CI 1.23 to 9.50) 5.20 (1.61 to 16.78) and 6.93 (1.55 to 30.99) respectively).[43] The Study of Osteoporotic Fractures (SOF), a prospective US study of 9704 ambulatory white women aged 65 and over, also found that a history of clinically-diagnosed VCF was strongly predictive of impaired function (age-adjusted OR 2.32, 95% CI 1.89 to 2.86).[41] Finally, Ryan found that 60% of women with symptomatic VCF attending a specialist bone clinic reported disturbed sleep; there was a significant association between sleep disturbance and the severity of vertebral deformities ($p<0.05$).[44] However, Ettinger et al found that women aged 55 to 75 with moderate to severe vertebral deformities were no more likely to require help at home because of their back than were similar women without vertebral deformities.[45]

Psychosocial outcomes

Ross has identified four categories of psychosocial problem associated with osteoporosis. These relate to:

- Quality of life
- Fears, anxiety and depression

- Self-esteem
- Social support and roles.[10]

However, he notes that these categories often overlap.

Quality of life

In people with osteoporotic fracture, quality of life can deteriorate quickly, even when physical function is not drastically affected, if changes in physical appearance, fear of fracture, and impediments to social function cause loss of self-esteem.[10] While most of the relevant research has been performed in postmenopausal women, men with VCFs and primary or secondary osteoporosis attending a UK hospital bone clinic scored much more highly in all six domains of the Nottingham Health Profile than age-matched or elderly male controls; the difference was particularly marked for energy, pain, and physical mobility. The physical mobility scores indicated greater disability in men with secondary osteoporosis than in those with primary osteoporosis ($p<0.05$).[46]

Fears, anxiety and depression

Symptomatic VCFs are associated with fears, anxiety, and depression which may relate to fear of future fractures, fear of loss of independence, and a feeling of hopelessness resulting from being told to avoid activities such as bending, twisting, and lifting heavy items, without being given advice on how to compensate.[10] Although postmenopausal women with a single VCF retain a good quality of life, once they have more than one fracture their quality of life is adversely affected by high levels of anxiety caused largely by fear of future fractures.[37] Such anxiety often leads to inactivity, which in turn can exacerbate bone mineral density loss and declines in physical fitness, thus increasing the risk of falling.[10,37]

In a small case-control study, Lyles et al found that psychiatric symptoms, as measured by the Hopkins Symptom Checklist Revised (SCL-90-R), were significantly worse in women with VCFs ($p=0.043$) than in matched controls; however, there was no significant difference in depression as measured by the Beck Depression Inventory ($p=0.129$).[38] Cook et al found that emotional problems were common in postmenopausal women with chronic back pain due to VCF: 82% reported fear of falling, while 66% reported frustration and 53% reported anger.[42] Unfortunately, this study did not include a control group of similar women without chronic back pain due to VCF.

Self-esteem

VCF may lead to height loss, spinal deformity, and abdominal protrusion, which adversely affect self-image and self-confidence, and to functional limitations which may lead to a loss

of self-esteem by limiting independence and the ability to participate in social activities.[37] Even relatively mild chronic pain may cause discomfort which discourages participation in social activities which involve sitting or standing for extended periods. Moreover, spinal curvature and height loss may make it difficult or impossible to sit or stand erect, causing problems with conversation and other activities.[10] Cook et al found that over 50% of postmenopausal women with a diagnosis of chronic back pain due to osteoporotic VCF reported problems with leisure/social activities.[42] However, in a small case-control study, Lyles et al found that women with VCFs and matched controls did not differ significantly in self-esteem as measured by the Rosenberg Self-Esteem Scale (p=0.731).[38]

Social support and social roles

The pain and physical impairment caused by VCFs can undermine the reciprocity involved in interpersonal relationships by reducing the ability to provide help and support to family and friends, while potentially increasing the need for assistance with activities of daily living and other personal care. If people are obliged to give up work, domestic, recreational, or sexual activities because of the limitations on their physical and functional abilities, they may also be deprived of their social roles.[37] The impact may be severe even if the activities in question do not seem to others to be demanding: inability to stand or sit for extended periods may limit involvement in social events, leading to an inability to fulfil the social roles which form an important source of self-esteem, and thus to a severe reduction in quality of life.[10]

- Significance for the NHS

Osteoporotic VCFs are associated with significant morbidity, mortality, and health and social care costs.[39,47] In a large UK-based study, Puffer et al.[48] found that, compared with matched controls, women diagnosed with osteoporotic VCFs had significantly more GP consultations (difference: 4.69, 95% CI: 4.35 to 5.03, p<0.001), referrals (difference: 0.51, 95% CI: 0.45 to 0.58, p<0.001), and hospital admissions (difference: 1.77, 95% CI: 1.63 to 1.91, p<0.001) in the year following diagnosis. The rate of GP consultations, referrals, and hospital admissions were also significantly higher in the year prior to diagnosis (all, p<0.001). Based on these figures, Puffer et al. estimated difference in costs per patient of £1015 and £1598 for pre- and post-diagnosis years, respectively.[48] Furthermore, it was found that patients with VCFs had a significantly greater utilisation of pharmacological treatments in the year following diagnosis, with the largest difference being in the prescription of bisphosphonates (difference: 52.71%, 95% CI: 49.37 to 56.01, p<0.001). The total additional cost of pharmacological treatment per patient was estimated to be £97.37 per year.[48]

Nevertheless, there are several reasons to interpret these estimates with caution. As noted above, people diagnosed with osteoporotic VCFs are more likely to have significant comorbidities requiring medical care. Hence, it is difficult to establish whether additional resource usage arises directly from the VCF. Furthermore, although only 30% of VCFs come to medical attention,[49] undiagnosed VCFs are also likely to be associated with greater service use due to the association of VCFs with excess morbidity and mortality.[50] The limitations VCFs can place on participation and consequently, on patient well-being, are highly significant issues with respect to care provision.[51]

- Measurement of disease

Osteoporotic VCF is identified by diagnosing both osteoporosis and vertebral fracture. The generally accepted approach to osteoporosis diagnosis is by the measurement of bone mineral density (BMD). The presence of osteoporosis is assessed by converting an individual patient's BMD into a measure known as the T-score, that is, the number of SDs from healthy young adults matched for ethnicity and gender. A T-score $\leq$ -2.5 is widely accepted as the diagnostic threshold.[52] A meta-analysis has shown that the predictive value of a 1 SD decrease in bone mass for osteoporotic fractures was roughly similar to that of a 1 SD increase in blood pressure for stroke, and more than a 1 SD increase in serum cholesterol concentration for cardiovascular disease.[53] Methods of assessing BMD include single-photon and x-ray absorptiometry (SPA and SXA) of the forearm and heel, dual x-ray absorptiometry (DXA), and dual photon absorptiometry (DPA) of the lumbar spine, proximal femur, whole body or particular regions thereof, and quantitative computed tomography (QCT) of the spine or appendicular sites.[2]

A number of methods have been proposed for identifying vertebral fractures. A widely used approach is the semiquantitative technique first described by Genant and colleagues,[7] which also indicates fracture severity. This approach utilises predefined thresholds for fracture severity, based on perceived reductions in vertebral height and area. Hence, vertebral bodies can be classed as normal (grade 0), mildly deformed (grade 1: reduction between 20%-25% in anterior, middle, and/or posterior height and a reduction of area of 10%-20%), moderately deformed (grade 2: reduction between 25%-40% in anterior, middle, and/or posterior height and a reduction of area of 20%-40%), and severely deformed ($\geq$40% reduction in any height and area). This grading system has demonstrated good to excellent intra- and inter-observer agreement, and similar estimates of incidence to quantitative morphometric measurements of vertebral height loss.[7] Common measures of angular deformity include kyphotic wedge angle, sagittal index, and measures of sagittal balance, in particular lateral radiographs measuring

the relationship between the C7 and S1 vertebrae (these are discussed in more detail in section 4.1).

3.2. Current service provision

- Management of disease

Traditionally, VCFs have been treated with optimal pain management (OPM). Bed rest is often required for one to two weeks, until the acute pain begins to subside, and therefore hospitalisation may be necessary.[15] Pain relief is generally achieved with oral analgesics: narcotics can be effective for fracture pain, while non-steroidal anti-inflammatory agents (NSAIDs) may relieve pain of inflammation and muscle spasm associated with VCF.[54] Calcitonin has also been shown to have a strong analgesic effect on patients with acute osteoporotic VCFs.[55] Patients who develop radicular pain due to compression of the nerve root may also require a nerve-root block or epidural injection of steroid and an anaesthetic. If such pain becomes chronic, other medications such as antidepressants, anticonvulsants, and alpha-2-agonists may be required.[54] External immobilisation (back bracing or casting) may also be used to reduce pain and promote appropriate posture, although this strategy carries the risk of muscle tone loss.[54] Anti-osteoporotic medication should be prescribed to reduce the risk of further vertebral fractures.[15,16]

In order to prevent further bone loss, mobilisation should begin as soon as the acute pain begins to subside, and spine extension exercises may be used to strengthen the back muscles.[15] Muscle spasms associated with acute VCFs may be treated with muscle relaxants and heat treatment; massage and physiotherapy may also be required by patients with kyphosis.[56] Patients should also receive walking aids and education about ways to avoid pain in activities of daily living.[54]

However, many patients complain of progressive pain and progressive functional limitation and loss of mobility despite conservative management. Thus, 75% of patients (n=107) who were admitted to a Swedish emergency unit with a painful acute VCF and received conservative treatment reported persistent back pain at 12 months.[57] Moreover, conservative management cannot prevent kyphotic deformity.[58]

In theory, open surgery with internal fixation may be performed in patients whose pain does not resolve with conservative management. However, such surgery is rarely undertaken in osteoporotic patients because the poor bone quality reduces the likelihood of achieving good results, whilst comorbidities in this patient group increase the risks associated with surgery.[59]

Consequently, open surgery is generally only performed in patients with neurological deficits,[60] in whom the balance of risks and benefits differs from that in patients without such deficits.

OPM is associated with an increased risk of complications of bed rest (e.g. pneumonia, deep vein thrombosis, and pulmonary embolism[61]), side-effects of medication, admissions to nursing home, and death.[62] Narcotic analgesics may lead to debilitating side effects, in particular cognitive impairment, nausea, and constipation, while NSAIDs are associated with gastrointestinal side effects such as nausea, gastritis, and ulcers.[54] Injected calcitonin may cause side effects such as nausea and flushing,[55] whereas nasal calcitonin is mainly associated with rhinitis and nasal symptoms.[63] Additional medications which may be used for chronic pain are also associated with a range of side effects.[54] Even in the absence of such severe adverse effects, extended bed rest, and the use of back bracing or casting, may be problematic for many older patients: bed rest may result in loss of bone density and muscle mass, and braces are often poorly tolerated.[19]

- Current service cost

Medtronic[34] reference Strom 2011[31] as estimating the cost of treatment of a vertebral fracture in the UK to be approximately 2,756 Euros in the first year. However, Strom gets this figure from Stevenson et al 2006.[64] Medtronic also reference Swedish data that the total cost is almost as high as the cost of treatment of a hip fracture, with lower initial hospital costs offset by higher community and informal care costs between 12 and 18 months (see p 18).[65]

Synthes state that HES data for the last 12 months (apparently for the UK rather than England and Wales) recorded that 6,375 patients (undifferentiated, i.e. not all osteoporotic) who had no surgical intervention (excluding facet injection or analgesia) occupied 78,923 bed days, with an average length of stay of 12.38 days; a further 698 patients received surgical treatment (PVP or BKP with or without stent), with an average length of stay of 7.5 days for PVP and 5.9 days for BKP.[32] In their submission, Medtronic indicated that the average inpatient stay associated with BKP was 5.1 days,[34] while Johnson & Johnson identified the average length of stay as 3.24 days for PVP and 4.48 days for BKP,[35] with these values provided by Dr Foster Intelligence. The longer lengths of stay identified by Medtronic include patients receiving vertebral augmentation for trauma or malignancy.[34]

The Assessment Group note that a recent review of the cost-effectiveness of vitamin K compared with alendronate used a cost of a vertebral fracture in the first year of £2981.[66] This

estimate was based on 2006 costs, which had been inflated by 8% to meet expected 2008 costs.

- Variation in services and/or uncertainty about best practice

There is no single standard of best practice care provision for people with osteoporotic VCFs because treatment needs can vary substantially according to age, BMD loss, mobility, and broader life conditions. Hence, care packages tailored to individual needs have been recommended by a number of authors.[67-69] However, the general aim of rehabilitation is to restore mobility, reduce pain, and minimise the incidence of new VCFs. Barriers to adequate treatment in older people with osteoporosis include polypharmacy, comorbidities, and cognitive impairment. Therefore, prevention of pain, disability, and functional decline should be pursued with these constraints in mind.[68]

Analgesic treatment varies according to pain severity and patient-level contraindications. The need for back pain relief can typically be met with acetaminophen, and supplementary codeine for breakthrough pain.[70,71] In cases of more severe and persistent pain, narcotic analgesics may be required for satisfactory pain reduction. While short-term use of these drugs is unlikely to lead to adverse events, undesirable side-effects, in particular delirium and constipation, tend to be more pronounced in frail older people.[68] NSAIDs are often prescribed to treat low back pain; however, these drugs have been linked to gastrointestinal side-effects. Chronic use of NSAIDs is also known to pose a risk of potentially fatal gastroduodenal bleeding.[72]

A number of physical approaches to pain relief may also be beneficial for people with osteoporotic VCFs, although their efficacy remains moot. Back bracing is often used to minimise postural flexion and paraspinal muscle spasm, and to facilitate bone healing.[73] While there is moderate evidence that lumbar supports are effective for the treatment of general back pain,[74] their effectiveness in osteoporotic VCFs remains poorly understood.[75] Moreover, chronic use of braces may lead to weakening of the paravertebral muscles and increased pain.[76] There is limited evidence that massage and superficial heat and cold therapy reduces general back pain, although evidence is lacking for the effectiveness of either treatment in osteoporotic VCFs specifically.[77,78]

Due to their role in skeletal homeostasis, calcium and vitamin D are widely viewed as the first line in osteoporosis treatment. However, while higher doses of vitamin D may be associated with greater benefits, this effect is yet to be confirmed.[79] Furthermore, a recent meta-analysis found that calcium supplements without co-prescribed vitamin D led to an increased risk of

myocardial infarction.[80] Until recently, hormone replacement therapy (HRT) was widely used to treat women with osteoporosis. A meta-analysis of RCTs comparing the risk of VCFs in women treated with HRT compared with placebo found a risk reduction of approximately 33% in the HRT cohort.[81] However, HRT has been linked to a number of adverse events (AEs) including a 2.3% increase in the relative risk of breast cancer, and an association with venous and pulmonary thromboembolism.[82] While these AEs are linked only with current or recent use, they nevertheless suggest that HRT should be prescribed with caution.[82]

One possibility explored in a recent nonrandomised cohort study[83] was the use of local anaesthetic and facet joint injection to control VCF-related back pain. Wilson et al performed facet joint injections under fluoroscopic guidance with lidocaine 1% and bupivacaine 0.5% to anaesthetise the affected area. Approximately a third of the treated cohort (21 of 61) responded well to the intervention, which led these investigators to hypothesise that facet joint injections may be effective among patients in whom pain does not arise directly from the VCF, but from biomechanical effects of the VCF occurring elsewhere in the spine. Anecdotally, the use of this approach prior to more invasive techniques is now widespread in the UK, and indeed, is explicitly recommended by some NHS trusts.[84] However, its long-term effectiveness is yet to be assessed.

- Relevant national guidelines, including National Service Frameworks

NICE has issued Interventional Procedure Guidelines (IPG) on the use of vertebroplasty and balloon kyphoplasty:

- NICE IPG 12,[85] issued in 2003, states that percutaneous vertebroplasty may be considered for the provision of pain relief in patients with severe painful osteoporosis with loss of height and/or compression fractures of the vertebral body only if their pain is refractory to more conservative treatment.[85]

- NICE IPG 166,[86] issued in 2006, states that balloon kyphoplasty may be considered in patients with vertebral compression fractures whose condition is refractory to medical therapy and in whom there is continued vertebral collapse and severe pain.

Both guidelines stipulate that the procedure should only be undertaken:

- by clinicians trained to an appropriate level of expertise
- following discussion by a specialist multidisciplinary team which includes a radiologist and a spinal surgeon

- where there are arrangements for good access to a spinal surgery service.

3.3. Description of technologies under assessment

- Summary of Intervention

Percutaneous vertebroplasty

Percutaneous vertebroplasty (PVP) is a procedure in which bone cement (such as PMMA, glass polymers, hydroxapetite or calcium compound) is injected into a fractured vertebra with the intention of reducing the pain caused by bone rubbing on bone and strengthening the bone so that it is unlikely to fracture further.[56] PVP is most commonly performed in the thoracic and lumbar vertebrae, and only occasionally in the cervical spine.[14] It is additional, rather than an alternative, to conventional therapy.[56]

PVP is performed under radiological guidance using fluoroscopy.[87,88] It is usually performed using conscious sedation and local anaesthesia of the skin, subcutaneous tissue, and the periosteum of the vertebral body into which the needle is to be introduced.[88] Sedation or light general anaesthesia is used in the majority of cases, with decisions being based on patient-level contraindications and anaesthetist preferences.[89] After adequate infiltration of local anaesthetic, a small skin incision is made, and a disposable bone biopsy needle or trocar needle is placed centrally in the vertebral body using an image guided safe access route. This may be done bilaterally through the pedicles, oblique across one pedicle or lateral oblique through the base of the pedicle. Under constant screening, it is advanced through the pedicle into the vertebral body with the aid of a light orthopaedic hammer.[14] An 11- or 13-gauge needle is used.[90] The cement is then injected very slowly, again under constant fluoroscopic screening, and the injection is stopped immediately if the cement begins to spread into a blood vessel or towards the posterior cortical margin.[18] To achieve optimal vertebral filling, two trocars may be used, one on either side of the midline.[89] The procedure may last from 45 minutes to an hour, depending on the number of vertebrae being treated.[14] Some centres perform CT scanning at the end of the procedure to assess the distribution of cement and identify any complications.[14]

At the end of the procedure, the patient remains on the operating table until the cement within the vertebral body has set.[16] This usually took about 20 minutes with earlier generations of bone cement.[91] However, setting time has been substantially reduced in recent years. For example, Goto et al compared the setting time of PMMA with bone cements containing micron-sized titania particles, and found a setting time of 11 minutes in a commercially available PMMA-based cement (Osteobond, Zimmer, Warsaw, IN, USA).[92] Glass-based polymers can set within two to three minutes.[93] The patient should then be kept in the

recumbent position, with monitoring of vital signs and neurological evaluations every 15 minutes for the first hour and then every 30 minutes for the next two hours.[16] The initial mobilisation should be supervised by qualified staff.[14]

PVP may be done as a day case if the patient's general health and social circumstances are appropriate.[89] However, in exceptional cases, an overnight stay may be required.[14] Prophylactic intravenous antibiotics may be used both before and after the procedure; some operators limit their use to patients with immunodeficiency.[14] Non-steroidal or steroidal anti-inflammatory drugs may be used for two to four days after vertebroplasty to minimise any inflammatory reaction to the heat generated by the polymerisation of the bone cement.[16] This is unlikely to apply to glass polymer-based bone cements, which do not have the same exothermic reaction upon mixing.

A number of bone cements are available for carrying out PVP, and decisions can be based on patient needs and operator preferences. The high viscosity CONFIDENCE SPINAL CEMENT SYSTEM™ is marketed by Johnson and Johnson, and carries an average cost of £1546 per operation (see also section 6.2). Low viscosity cements are also available to purchase at prices that are lower than that of high-viscosity PMMA cement. The list price for such cements were obtained through NICE, and on clinical advice it was estimated that the costs using lower-viscosity cements, incorporating injection kit, needles cement and assorted consumables would be in the region of £660, £720 and £780 for one-, two- and three-level procedures respectively. However, our clinical expert estimated that 15% of cases are more complex and would require Cortoss® cement, collation or thicker cement, whilst younger patients would need bone absorbable cement. It was assumed that the added cost of these complex cases would add slightly over £100 to the average cost of an operation resulting in an assumed cost of £800 per low-viscosity cement PVP procedure. Given that the estimate includes a component for using higher viscosity cement, the price used within the analysis could be equated to a strategy where low-viscosity cement is used within the majority of patients, whilst higher-viscosity cements are used in a small proportion where the clinician believes that this is appropriate. In addition to the cements themselves, operating equipment, including bone biopsy or special trocar needles and vacuum cement mixing systems, are required.

Percutaneous balloon kyphoplasty

Percutaneous balloon kyphoplasty (PBK) is a variant of PVP in which one or two balloon-like devices (also known as tamps) are inserted bilaterally into the vertebral body. These balloons are slowly inflated until they reach their highest achievable volume, in order to restore

vertebral body height. The balloons are then deflated and removed, leaving a cavity which is filled with bone cement; because of the existence of the cavity, the cement may be injected at a lower pressure than is used for PVP.[34]

Sedation is based on practical considerations, such as patient-level contraindications. Some patients receive general anaesthetic and remain in hospital overnight for observation.[90] However, BKP may be done as a day case if the patient's general health and social circumstances are appropriate.[89]

Although there is no apparent reason why BKP should differ from PVP in terms of pain relief, it has some potential additional benefits. Medtronic suggest that the creation of a cavity of known volume into which cement may be injected results in a lower risk of leakage and consequent complications.[34] Furthermore, introduction of the balloon provides the potential for restoring vertebral height and thus correcting deformity. However, neither of these potential benefits have a good evidence base. There is no evidence of a higher complication rate in PVP, as most cement leakages remain asymptomatic. In addition, to be effective in restoring vertebral height and reducing kyphosis, BKP should be performed within a few weeks of fracture: thereafter, although a cavity will still be created within the vertebra, fracture healing is likely to prevent restoration of vertebral height.[94]

In the UK, the device required for BKP is marketed by Medtronic as a single-use sterile pack containing two Kyphon® Xpander™ inflatable bone tamps and associated accessories, at a list price of £2600.50.[34] The components of the Medtronic KYPHON BKP kit obtained the CE Mark in May 1999, while various components of the Kyphopak – the Osteo Introducers, Balloon, and Bone Fillers – obtained the CE Mark in 2001, 2003, and 2002 respectively.[34] The bone cement included in the kit, Kyphon® ActivOs™ Bone Cement with Hydroxyapatite,[95] is a PMMA cement to which hydroxyapatite (a calcium compound believed to promote osseointegration) has been added.[96] Kyphon® also produces two alternative cements for use in kyphoplasty: Kyphon® KyphOs FS™ Calcium Phosphate Bone Substitute and Kyphon® HV-R™ Bone Cement.[95]

Balloon kyphoplasty with stenting (stentoplasty)

BKP with stenting seeks to overcome a problem inherent in simple PBK, namely that, because of pressure on the vertebra, some of the height restored by the fully inflated balloons may be lost after they are deflated and removed and before the cement is injected. A laboratory comparison of stenting compared with kyphoplasty found that most of the height gained in balloon kyphoplasty appeared to be lost after the balloon was deflated.[97] In

stentoplasty, a small balloon catheter surrounded by a metal stent is inserted into the vertebral body using a minimally invasive percutaneous approach under radiographic guidance and either local or general anaesthetic. The balloon catheter is then inflated with liquid, under pressure, to create a cavity in which the stent is expanded. The balloon catheter is then deflated and withdrawn but the stent is left in position within the vertebra and maintains the height of the cavity into which high-viscosity PMMA bone cement is then injected. The injected cement hardens within an hour, and the patient may then be mobilised.[32]

The use of a vertebral body balloon (VBB), an optional site preparation device, is recommended: it enables the operator to identify the feasibility of cavity creation and full expansion of the stents.[32]

Synthes market a vertebral body stenting (VBS) system which consists of a vertebral body stent catheter, an inflation system, a VBS access kit, and an optional vertebral body balloon (VBB) catheter. The balloons included in the VBS and VBB catheters are said to be considerably more rigid than current kyphoplasty balloons, and therefore less likely to herniate through the fracture.[32] Anecdotally, stenting is associated with a greater risk of procedure-generated adjacent fractures, and some operators cement the adjacent vertebrae as a preventive measure.

Facet joint injection

Facet joint injections involve the administration of anti-inflammatory steroids and local anaesthetic to facet joints with focal tenderness. They are usually performed on an outpatient basis without sedation. Prior to the procedure, the patient lies in prone position, while the operator palpates the back in order to localise the pain. Once identified, the skin and subcutaneous tissue surrounding the affected area are infiltrated with 1% lidocaine. Then, fluoroscopic or CT imaging is used to identify the posterior part of the joint, and a 20- or 22 gauge needle is directed vertically into the joint space until bone cartilage is reached.[98] A long-acting steroid such as triamcinolone, methylprednisolone, or betamethasone is administered to the joint, along with 0.5% bupivacaine to anaesthetise the area.[99]

Although we did not directly assess the clinical effectiveness of facet joint injections for treating painful osteoporotic VCFs, the procedure is noted here because it is emerging as a possible treatment for a subgroup of patients in whom pain and functional impairment arises not from the VCF per se, but from the impact of the VCF on other spinous processes. Ryan et al[100] found that facet joints may be an important site of pain for people with osteoporotic VCFs, and more recently, a cohort study by Wilson et al (n=61 treated patients)[83] found that

problems with the facet joint may account for back pain in approximately one third of patients with osteoporotic VCFs. On the basis of their data, Wilson et al. suggested that the apparently high placebo response seen in the 'sham' treated cohorts in two recent RCTs[101,102] may in fact have been a response to the local anaesthetic by patients whose pain was an indirect consequence of a VCF. Hence, in our cost-effectiveness model, a hypothetical scenario controlling for the potential influence of this patient subgroup has been included.

Bone cement

PVP and BKP are traditionally performed using polymethylmethacrylate (PMMA), a low-viscosity acrylic bone cement, to which a radio-opaque substance such as barium, tantalium or tungsten sulphate has been added to facilitate visualisation during the procedure.[14] It is prepared by mixing a liquid component containing the monomer, accelerator, and inhibitor, with a powder containing the polymer, radio-opacifier, and initiator. The heat which is released during the subsequent polymerisation process while the cement is hardening in situ may cause local damage to bone or other tissues.[103] However, in PVP and BKP, such damage may not be entirely detrimental. PMMA appears to have analgesic properties quite apart from those caused by the effect of the stability provided by the cement within the weakened vertebrae. The reason for such analgesic properties remains unclear, but one possibility is that it destroys or damages local nerve endings as a result of both the toxic effects of the free monomers of PMMA and the heat caused by the cement polymerisation.[14] However, we are not aware of any evidence to suggest that cements that do not generate heat are any less effective.

The FDA[103] state that PMMA is contraindicated in the presence of active or incompletely treated infection at the site where the cement is to be applied. They also note that hypotensive reactions have been noted between 10 and 165 seconds after its application; as these have lasted from 30 seconds to over 5 minutes, and some have progressed to cardiac arrest, they recommend that patients should be monitored carefully for any changes in blood pressure during and immediately following the application of the cement. Other reported adverse events include pyrexia due to allergy to the cement. In addition, they note that the heat released while the cement is hardening in situ may damage bone or other tissues surrounding the implant.[104]

The FDA also note that caution is required in preparing and handling PMMA: excessive exposure to the concentrated monomer vapours may produce irritation of the respiratory tract, eyes, and possibly the liver; contact lens wearers should not be near or involved in mixing

PMMA.[103] However, newer manufacturer kits, such as the PLACOS® bone cement (Zimmer, Hanau, Germany) provide vacuum cement mixing tools to avoid this issue.

The newer composite cement bisphenol-a-glycidyl dimethacrylate (bis-GMA) resin (Cortoss) is more viscous than PMMA, and consequently easier to handle. It does not contain the volatile monomers which may be the cause of cardiovascular and respiratory AEs with PMMA. It is stronger than PMMA, and cures at a lower temperature, reducing the risk of thermal damage, and setting more rapidly. It is also inherently opaque, and therefore does not need to be mixed with a toxic radio-opaque material.[14] The bioactive Cerament™ bone analogue cement (Bonesupport, Lund, Sweden) also has radio-opaque properties which obviates this requirement.

Follow-up required

It has been suggested that, following discharge, patients who have undergone percutaneous vertebroplasty should be recalled for evaluation one day, one week, and one month after treatment.[91] However, this appears to reflect US practice, and according to our clinical expert (DW) in the UK it would be more normal for a patient to receive a follow-up telephone call at one week after discharge, and a clinical consultation at one month.

Setting

PVP and BKP should be performed in a sterile environment which allows fluoroscopic imaging of the thoracolumbar spine.[18] The use of an interventional radiology suite rather than an operating theatre has been recommended because fixed fluoroscopic equipment offers better imaging quality than a mobile C-arm.[16] PVP and BKP should only be performed in hospitals which have adequate neurosurgical backup to deal with potentially serious complications.[14]

Equipment required

Magnetic resonance imaging (MRI) equipment is requisite to screen all patients who are considered for PVP or BKP, in order to identify the fracture, assess its age, define its anatomy, assess the posterior vertebral body wall, and exclude other causes of back pain. However, Computed tomography (CT) scanning may be used instead when MRI is unsafe (for example, in patients with pacemakers). CT equipment is also required if there are any doubts regarding the intactness of the posterior vertebral wall.[16] Fluoroscopic imaging equipment is also required for use during the procedure.

Personnel involved

As stipulated in the NICE guidance,[85,86] PVP and BKP should only be performed by clinicians trained to an appropriate level of expertise (for historic reasons, PVP and BKP have most commonly been performed by interventional radiologists). An anaesthetist should preferably be present to monitor sedation even when the procedure is performed under local anaesthesia.[14]

Place in the treatment pathway

PVP and BKP are usually offered as a last resort to people with symptomatic vertebral compression fractures in whom alternative treatments have not been successful.[88,105] An Australian review has noted that, while people with recent painful VCFs are potential candidates for either PVP or BKP, PVP is not appropriate for people with VCFs which cause symptoms such as pain or breathlessness due to a hunched posture, and who require structural correction for functional kyphotic deformity which is neither congenital nor due to trauma: such patients are potential candidates for either BKP or surgical stabilisation of the fracture with or without fusion of the vertebrae.[56]

Criteria for treatment

NICE guidance indicates that PVP and BKP should be limited to patients whose pain is refractory to more conservative treatment;[85,86] for BKP, there is an additional requirement that they should have continued vertebral collapse and severe pain.[86]

Recent guidance from the Cardiovascular and Interventional Radiology Society of Europe (CIRSE) states that PVP is indicated in patients with "painful osteoporotic VCF refractory to medical treatment. It defines failure of medical treatment as "minimal or no pain relief with the administration of physician-prescribed analgesics for 3 weeks or achievement of adequate pain relief with only narcotic dosages that induce excessive intolerable sedation, confusion, or constipation".[16] The CIRSE guidelines further note that PVP may be considered within days of painful fracture if the patient is at high risk of complications resulting from immobility (e.g. thrombophlebitis, DVT, pneumonia, or pressure ulcer).[16]

Contraindications

The CIRSE guidelines list the following absolute contraindications to PVP:

- Asymptomatic vertebral body compression fracture
- Patient improving on medical treatment
- Osteomyelitis, discitis, or active systemic infection

- Uncorrectable coagulopathy
- Allergy to bone cement or opacification agents
- Prophylaxis in osteoporotic patients.[16]

Relative contraindications in osteoporotic patients include:

- Radicular pain
- Fracture of the posterior column (which increases the risk of cement leak)
- Spinal canal stenosis
- Lack of surgical backup and monitoring facilities.[16]

These contraindications appear to be equally applicable to BKP.

Although neurological symptoms are not an absolute contraindication to PVP or BKP, in patients with such symptoms great care should be taken to avoid cement extravasation as this may exacerbate any pre-existing nerve compression.[89] Thus, prior to PVP/KP, a detailed examination should be performed to detect any neurological compromise and exclude other causes of pain such as degenerative spondylosis. Physical examination is important to accurately localise the symptomatic vertebra, especially in the presence of multiple fractures.[14]

- Identification of important sub-groups

A number of authors have suggested that only patients with acute VCFs (≤6 weeks' duration) are likely to benefit from PVP or BKP.[106,107] However, the clinical experience of rapid and dramatic post procedural reductions in pain is likely to be confounded by the rapid healing of the fracture itself. This was suggested by the recruitment pattern in the recent VERTOS II trial, in which more than half of initially eligible participants became ineligible prior to enrolment due to spontaneous pain reduction.[17] Furthermore, Rad and Kallmes[108] presented a retrospective analysis of 321 single-level PVP procedures. These were stratified into acute (≤6 weeks), subacute (6 to 24 weeks), and chronic (>24 weeks) fractures, and absolute and proportional pain reductions were compared between the three groups. There was no strong correlation between fracture acuity and pain relief. Hence, vertebral augmentation may be better used to treat people with chronic pain refractory to more conservative measures.[84,109]

It has also been suggested that PVP is effective in specifically selected patients with more severe pain.[107] However, the evidence for this claim is unconvincing. Pain is a subjective experience mediated by various psychosocial factors[110] and is consequently open to

confounding influences and difficult to objectively assess. A strong placebo response due to positive expectations among persons with severe pain could not therefore be ruled out as a cause of apparent effectiveness. Indeed, a recent meta-analysis of IPD from two placebo-controlled trials of PVP[109] found that post procedural pain reduction was unrelated to baseline pain severity.

It is possible that vertebral augmentation should only be pursued with those whose pain and functional impairments arise directly from the VCF, rather than indirect consequences thereof. Wilson et al[83] found that a third of patients who were eligible for PVP (n=21 of 61 treated patients) responded favourably to a facet joint injection. It may be that the pain in this group was mediated by overload on the facet joints adjoining the fractured vertebral body, which could therefore be treated successfully with a less invasive procedure.

- Current usage in the NHS

On the basis of data from Dr Foster Intelligence, Johnson & Johnson have estimated that, between April 2010 and March 2011, 473 vertebroplasties and 225 kyphoplasties were performed for osteoporotic VCF.[35] These figures appear to apply to England alone.
Medtronic reported, on the basis of HES data for 2009/10/11, that 487 patients in England and Wales were treated with PVP and 466 with BKP, for osteoporotic VCF;[34] it is not clear whether this is figure relates to the two-year period or to the average for one year.

Synthes note that HES data for 2010/11 indicate that 698 patients in the UK underwent either PVP or BKP with or without stent for osteoporotic VCF.[32]

Anticipated costs associated with the intervention

A formal report on the likely costs associated with each analysis is contained in the cost-effectiveness section.

4. DEFINITION OF THE DECISION PROBLEM

4.1 Decision problem

The assessment will address the question "What is the long-term efficacy, safety, and cost-effectiveness of percutaneous vertebroplasty and percutaneous balloon kyphoplasty (with or without vertebral body stenting) as a treatment for osteoporotic vertebral compression fractures?"

Interventions

Percutaneous vertebroplasty (PVP) or percutaneous balloon kyphoplasty (BKP) with or without vertebral body stenting, performed under general or local anaesthesia.

Population including sub-groups

The relevant population is adults of any age and either gender with painful osteoporotic vertebral compression fractures. If the evidence permits, consideration will be given to subgroups defined by:

- time from fracture to treatment
- presence of fracture-related deformity before treatment
- receipt of inpatient care before treatment.

People with malignancy-related vertebral fractures, and those with neuropathy in the absence of osteoporotic compression fractures, are not included the scope of this assessment.

Relevant comparators

The comparators specified in the scope are the interventions themselves, and non-invasive management (including no treatment in people who cannot tolerate the relevant active comparator interventions). Injection of local anaesthesia to the affected vertebral body is also considered a relevant comparator, as this has been used as a 'sham' intervention in double-blind, placebo-controlled trials of PVP. Moreover, our clinical advisor (DW) suggested that administration of local anaesthesia with facet joint injection is now routinely offered in the UK as a minimally invasive intervention before considering patients for vertebral augmentation.

Both the Buchbinder[101] and INVEST[102] studies used what they describe as a 'sham' intervention for the control procedure. The procedure in each of these trials involved infusion of lidocaine 1% into the skin to numb the affected area. The INVEST trial also infiltrated the

periosteum of the pedicles with 0.25% bupivacaine. Both trials then mimicked vertebroplasty through physical cues such as pressure to the back, and opening of the methacrylate monomer to simulate the smell associated with PMMA preparation. In this review, it was decided that, rather than sham, these procedures would be described as 'operative placebo with local anaesthesia' (OPLA). This term was chosen because of the ongoing debate as to whether these procedures actually constitute a sham intervention. A number of authors have argued that the local anaesthetic may have had specific mechanisms of efficacy for long-term pain reduction,[106,111,112] and indeed, some empirical evidence is available to support this possibility.[83,113] Conversely, some practitioners have proposed that, due to the relatively low volumes of cement used in the Buchbinder and INVEST trials, the comparison was effectively placebo versus placebo.[114]

Therefore, it was also viewed as important to highlight the possibility of a high placebo response to these interventions, which could be much greater than the response associated with vertebral augmentation. This may be strongly influenced by the elaborate rituals required in any operative procedure. According to Kaptchuk, healing rituals comprise "compelling multi-sensory dramas involving evocation, enactment, embodiment and evaluation".[115] In surgical procedures, these rituals encompass the interventionist's language and dress, the hospital setting in which they are performed, and the lived experience of being anaesthetised and undergoing the intervention. In short, each dimension of the surgical ritual implies a scientifically derived and culturally sanctioned process designed to move the patient from an 'ill' to a 'well' state. Such rituals enhance suggestibility and so heighten the probability of a favourable outcome.[116] Consequently, it has been argued that researchers must take these suggestive effects into account, particularly when considering trials that measure subjective outcomes such as pain.[116]

There is no gold standard for non-invasive management: the American Academy of Orthopaedic Surgeons considers the strength of the evidence for the various non-invasive treatment options (such as physiotherapy, analgesia, and the use of anti-osteoporotic agents such as a bisphosphonate or strontium ranelate) to be generally weak to inconclusive, although they provide a recommendation of moderate strength for the short-term use of calcitonin.[117]

Outcomes

Primary outcomes

- Health-related quality of life
- Back-specific functional status/mobility

- Pain/analgesic use
- Vertebral body height and angular deformity
- Incidence of new vertebral fractures
- Progression of treated fracture

Secondary outcomes

- All-cause mortality
- Symptomatic and asymptomatic leakage of cement (e.g. into adjacent intervertebral discs)
- Periprocedural balloon rupture
- Post-operative complications (including infection)
- Other adverse events

The majority of the primary outcomes take the form of continuous or quasi-continuous outcomes (e.g. pain measured on a 0-10 scale), whereas the secondary clinical outcomes are binary outcomes (e.g. the number of patients who suffer a given complication). Continuous outcomes can be compared in terms of:

- The difference between the mean scores in the intervention and control groups at a specified point in time
- The difference between the change in mean score in each group between two specified points in time (e.g. immediately before and one month after treatment.

To ensure that a continuous outcome measures a real difference between the intervention and control groups following the intervention, either the pre-intervention score for that outcome must have been identical in both groups or any pre-intervention differences must be minimised or controlled for through statistical adjustment. For this reason, we have presented continuous outcomes in terms of changes from baseline rather than as results at specified points in time.

While it is easy to determine whether any of these differences in outcome are statistically significant, it is less apparent whether they are also clinically relevant – in other words, whether they represent differences which the patients would recognise as beneficial. For this reason, research has been conducted for some outcome measures to attempt to quantify the smallest change in score which reflects a change in symptom which can be considered clinically relevant: this is termed the minimal clinically important difference (MCID). It should be noted that the proposed MCID values were for individual rather than group changes: in a trial, individual patients may show clinically important improvements even

though the between-group difference in means is less than the MCID.[118] Nonetheless, the MCID provides a useful means of assessing whether, on average, an intervention is likely to be associated with greater clinical benefit than the control treatment.

Discussion of individual outcome measures

- Health-related quality of life

Generic measures of health status are particularly important in populations with comorbidities, such as the elderly, because disabilities from these comorbidities may influence the patient's response to treatment. Also, because such measures include mental and social health, they give a more complete picture of the patient's health than do back-specific instruments.[119]

The **Assessment of Quality of Life (AQoL)** measure was designed for use in the evaluation of health care interventions, and is sensitive to changes in the frail elderly. Its five scales measure illness, independent living, social relationships, physical senses, and psychological well-being, and each dimension may be reported separately. Alternatively, a single utility score may be computed by combining all of the scales except the illness scale; the score ranges from 1.00 (representing full HRQoL) to -0.04 (representing HRQoL states worse than death), with 0.00 representing death. The MCID appears to be 0.06.[120]

The **Dallas Pain Questionnaire (DPQ)** is a 16-item instrument designed to evaluate the effect of chronic pain on four aspects of life: daily activities; work and leisure activities; anxiety/depression; and social interest. Each item is scored on a scale whose extremities are marked 0% and 100%; although this scale appears to be continuous, for each item it is divided into five, six, seven, or eight segments, each of which has a score. The item scores for each of the four aspects are added together and multiplied by a constant. They are then reported as a percentage, with lower scores indicating better quality of life.[121] No MCID was identified for the DPQ.

The **EQ-5D** has five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) each assessed by a single question on a three-point ordinal scale (no problems, some problems, extreme problems). These are responses are combined and presented as a quasi-continuous outcome on a scale of -0.59 to 1.00, where 0 represents death and 1.00 indicates 'full health'; negative scores represent health states valued as worse than death. The estimated MCID for people with back pain is 0.08.[122]

The **Mini Mental State Examination (MMSE)** is designed to measure cognitive status in adults. It is scored from 0 to 30, with higher scores indicating better cognitive condition.[123]

The **SF-36** (Medical Outcome Study Short Form 36) is a questionnaire designed specifically to measure self-reported health-related quality of life; it has been proposed as the most appropriate measure of generic health status for use in people with spinal disorders.[119] It contains 36 questions which measure functional status, wellbeing, and overall health in eight dimensions (physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional health, and mental health); these eight dimensions may be aggregated to produce summary measures of physical health (the physical component score - PCS) and mental health (the mental component score - MCS). Results are presented on a scale of 0 to 100, with higher scores indicating better health, and may be transformed to take into account population norms.[124] Copay et al have suggested an MCID of 4.9 points specifically for the PCS,[125] whereas Angst et al have suggested an MCID for improvement of 2.0 in the PCS, 7.8 in the bodily pain subscale, and 3.3 in the physical function subscale.[126] Wiebe et al have suggested an MCID of 3.0 for the PCS and 4.6 for the MCS.[127] No MCID has been identified for the overall SF-36 utility score.

Some of the studies included in this review only used the PCS, ignoring the scales for vitality, social functioning, role limitations due to emotional health, and mental health, and thus undermining the value of using a generic measure of quality of life.

In addition to the generic quality of life measures, one measure, the **QUALEFFO** (Quality of Life Questionnaire of the European Foundation for Osteoporosis), has been developed and validated specifically for use in clinical trials in patients with vertebral osteoporosis. The **QUALEFFO-41** has 41 questions relating to five domains: pain, physical function, social function, general health, and mental health.[128] The domain scores are presented on a scale of 0 to 100 where 0 corresponds to the best HRQoL and 100 to the worst HRQoL.[129] QUALEFFO-41 has been shown to discriminate better between patients with and without vertebral fractures than the EQ-5D.[129] No MCID is known to have been suggested for the QUALEFFO. There is also a shorter version, the QUALEFFO-31.[130]

- Back-specific functional status/mobility

Measures of functional status assess the ability to perform specific tasks: this ability may only have a weak relationship with the reported level of pain.[131] However, for many patients, the greatest problem caused by a VCF is the limitation of activity rather than pain per se, and

therefore functional status is a more clinically meaningful outcome than pain status. Moreover, it has been argued that the Roland-Morris Disability Questionnaire (RDQ), one of the most commonly used back-specific scales, is more useful than self-reported pain for assessing the impact of vertebral fractures on the patient's daily life because it is more objective: self-reported pain is influenced by the patient's perception and tolerance of pain, and relate only loosely to functional limitation.[132] This argument also applies to other measures of disability.

Furlan et al differentiate between back-specific functional status, as measured by items such as the RDQ, and disability, which can be assessed in terms of factors such as the ability to perform activities of daily living, work absenteeism etc; they consider both to be important patient-centred outcomes, along with symptoms (e.g. pain), perception of overall improvement, satisfaction with treatment, and well-being (e.g. quality of life measured with the SF-36 etc).[133] The studies included in this review use a number of different measures of functional status.

The **Roland-Morris Disability Questionnaire (RDQ)** was designed to assess physical disability due to low back pain.[134] It assesses self-reported functional status in eight dimensions (physical activities, housework, mobility, dressing, getting help, appetite, irritability, and pain) by measuring 24 activity limitations.[135] Scores range from 0 (no disability) to 24 (maximum disability).[134] The original RDQ contained a 6-point pain rating scale. However, this has now been excluded with the recommendation that the SF-36 pain scale be used for this purpose.[134] The modified RDQ (RDQ-23) contains 23 questions, some of which differ from those in the original questionnaire; it is scored from 0 to 23.[121]

The MCID for the RDQ varies according to the level of disability of the patients, from 1-2 points in patients with little disability to 7-8 points in patients reporting high levels of disability, and 5 points in uncategorised patients; Ostelo et al suggest 5 points, or 30% improvement from baseline.[118] Lauridsen et al suggest an overall MCID of 5 points, but recommend that 2 points should be used in populations attending hospital back pain clinics.[136] However, Roland and Fairbank recommend that, in clinical trials which use the modified 23-item version of the RDQ, an MCID of 2-3 points should be used for sample size calculations.[134]

The **Oswestry Disability Index (ODI)** was designed to measure patient-reported disruption to activities of daily living attributed to back pain . It comprises 10 dimensions (pain intensity, personal care, lifting, walking, sitting, standing, sleeping, sex life (if applicable), social life,

and travelling), each containing 6 response categories scored from 0 (indicating no disruption) to 5 (the worst state). The score is reported as a percentage of the maximum possible score (i.e. on a scale of 0-100). The weighted mean score for "normal" populations is 10.19.[134]

It has been suggested that, for the ODI, the MCID should be 4 points in relation to mean scores between groups but 15 points in patients before and after surgery.[134] More recently, Ostelo et al have suggested an MCID of 10 points in the transformed (0-100) scale, or 30% improvement from baseline,[118] while Lauridsen et al suggest an MCID of 11 points in all patients, and 8 points in patients attending hospital back pain clinics.[136]

Both the ODI and the RDQ are easy to use, reliable, and valid. Because of floor and ceiling effects, the ODI is recommended for use with patients who are likely to have persistent severe disability, and the RDQ with patients who are likely to have relatively little disability, but for most patient groups both instruments appear to function satisfactorily in groups with severe disability.[134] However, the modified, 23-item, version of the RDQ[137] has been shown to be more responsive than the ODI in patients with low back pain only, whereas the ODI appeared slightly more responsive in patients with leg pain and/or low back pain.[136] Moreover, self-reported disability measures such as the RDQ have been shown to display only modest correlation with direct measures of physical function.[134]

The **Study of Osteoporotic Fractures-Activities of Daily Living (SOF-ADL)** scale consists of 6 questions relating to activities undertaken in a typical day. The scale ranges from 0 to 18, with higher scores indicating more back-related disability.[123]

The **Barthel Index** is a 10-item scale designed to evaluate the observer-assessed ability of a patient with a neuromuscular or musculoskeletal disorder to care for him or herself. The items assess independence in key activities of daily living (feeding, transferring from wheelchair to bed and back, grooming, toilet use, bathing, mobility on a level surface, ascending and descending stairs, dressing, faecal continence, and urinary continence). It was originally scored from 0 to 100, with lower scores indicating greater disability,[138] but may also be scored from 0-20, and in either case lower scores indicate greater disability.

A change of 1 point in the 0-20 Barthel Index (5 points in the original 0-100 version) represents a change in level of dependency in any of the key activities, and is therefore likely to be clinically meaningful.[56]

Direct observer-assessed measures of physical function include:

- The **timed 'Up and Go' test**, which assesses functional mobility by measuring the time in seconds required to rise from a standard armchair, walk 3 metres, turn round, return to the chair, and sit down again.[139]
- The **tandem test**, which assesses balance by measuring the time for which the patient can stand in three different positions.[140]
- The **repeated chair test**, which tests muscle power by asking the patient to rise from, and return to, sitting as many times as possible in 30 seconds; higher scores indicate better health status.[123]

The included studies also assessed functional status in terms of the use of aids and appliances, days of bed rest or reduced activity, and perceived recovery.

- Pain/analgesic use

The recommended measure of global pain severity is the **bodily pain subscale of the SF-36**, a two-item scale which measures pain intensity on six levels (from none to very severe) and interference with activities on five levels (from not at all to extremely).[119] This measure is recommended not least because of the availability of normative data.[141] An absolute cut-off value has been suggested for the MCID of 3 points overall, or 2 points in populations attending hospital back pain clinics.[136]

However, many of the included studies used either:

- A **visual analogue scale** (VAS), whereby patients mark the point which best represents their pain on a line, usually 10 cm long, whose ends bear labels describing the extremes of pain intensity (e.g. 'no pain' and 'worst imaginable pain'); or
- A **numeric rating scale**, whereby patients rate the intensity of their pain on a scale of 0 to 10 (11-point scale) or 0 to 100 (101-point scale), where 0 represents no pain, and 10 or 100 the worst possible pain.[141]

The VAS, which is formatted without numbers, represents a continuous range of values, whereas the numeric rating scale is formatted using whole numbers to form a segmented scale.[142]

Because it is usually measured in millimetres, and can therefore be regarded as having 101 response categories, a 10-cm VAS is potentially more sensitive than an 11-point numeric rating scale. However, it has been shown to be more difficult to understand than other measures of pain intensity, especially for people at risk of cognitive difficulties (e.g. some

elderly individuals, or people taking high doses of opioid analgesics).[141] Moreover, the VAS was developed for assessing chronic pain and is less reliable in the immediate postoperative period, when any single VAS score may have an imprecision of ±20 mm (20%).[143] Numeric rating scales are easier to use, and their sensitivity and validity has also been demonstrated. Both Bolton and Wilkinson[144] and Grotle et al[145] suggest that the numeric rating scale may be more responsive than the VAS. However, unlike the VAS, numeric rating scales have not been demonstrated to have ratio qualities (i.e. a change in pain score from 9.0 to 6.0 cannot be assumed to represent a 33% decrease in perceived pain).[141]

Some of the included studies (Blasco,[146] Farrokhi,[147] Liu,[148] Rousing,[139] and VERTOS II[149]) stated that they used a VAS scale; others (Buchbinder,[150] FREE,[151] INVEST,[152] VERTOS[153]) specified that they used an 11-point numeric rating scale, although in the FREE[151] and VERTOS[153] studies this was called a VAS.

As the distribution of scores from VAS and numeric rating scales is not normal, non-parametric statistical analyses are appropriate.[56] Moreover, because initial and subsequent pain ratings on the VAS tend to be correlated, any between-group comparison should compare changes in scores from baseline rather than simply differences in the final post-treatment score.[56]

Ostelo et al have proposed an absolute cut-off value for the MCID of 15 on a 100 unit VAS, with a relative cut-off value a 30% improvement from baseline.[118] However, DeLoach et al suggest that, because of the imprecision found in the immediate post-operative period, the MCID in that period should be 20/100 units.[143] Ostelo et al also proposed that, when an 11-point numeric rating scale is used for low back pain, the absolute cut-off value for the MCID should be 2 points, again with a relative cut-off value of a 30% improvement from baseline.[118] However, Copay et al suggest a MCID of 1.2 points for back pain.[125]

The INVEST study also used a modified version of the Deyo-Patrick pain frequency and bothersomeness scale.[152] The original scale measured the frequency with which patients experience pain, and how pain impacts on their daily life, on a scale of 0 to 6, where 6 indicates the highest impact.[137] The INVEST study used a scale of 0 to 4, with higher scores again indicating more severe pain.[102]

While the technology assessment protocol did not specify analgesic use as an outcome, it has been included because some of the included studies used analgesic use as a proxy for pain relief. This may be measured either quantitatively (i.e. amount of analgesia used) or

qualitatively (i.e. type of analgesia used). Most of the included studies used a qualitative approach, grouping analgesics into categories such as non-opioid, weak opioid, and strong opioid. Doidge et al note that these distinctions are clinically important because the risk profiles of the different categories of analgesic differ substantially.[56]

- Vertebral body height and angular deformity

Vertebral body height

Vertebral body height may be measured at different parts of the vertebra (posterior, middle, and anterior). In addition, restoration of vertebral body height may be reported in four different ways:

1. Absolute restoration in millimetres
2. Percent restoration relative to preoperative height of the fractured vertebra
3. Percent restoration relative to lost vertebral height (based either on a pre-fracture x-ray or on an estimate of the unfractured height of the fractured vertebra)
4. Percent restoration relative to referent vertebral height (the height of the nearest nonfractured vertebra).McKiernan et al[154]

McKiernan et al found substantial variations in the reported magnitude of height restoration when x-rays of the same vertebrae were measured using all four methods. Unless the same referent normative height was included in each x-ray, the comparison of "absolute" values, both between studies and between x-rays in the same study, was unreliable because of the possibility of magnification error. The use of relative data allowed comparison both within and between studies which used the same fixed dimension, provided that the precision error of the measurement was acceptable. However, studies which used different fixed dimensions could not be compared: thus, for instance, the apparent magnitude of height restoration was almost four times greater when anterior vertebral height was measured using method 2 than when using method 4. The choice of fixed dimension may have a differential effect on fractures of varying degrees of severity: in severely compressed vertebrae, the apparent restoration was greater using the preoperative height (method 2) than the lost vertebral height (method 3), whereas, in very mild fractures, the restoration was greater using method 3 than method 2.[154]

McKiernan et al therefore recommended that reports of vertebral height restoration should:

- Include all index vertebral height dimensions (posterior, middle, and anterior)
- Include absolute measurements of all referent vertebral heights

- Be reported relative to a referent normative height (either referent vertebral height or a radio-opaque object of known dimension included in the radiographic field), in order to permit correction for inter-radiographic measurement error
- Take into account the dynamic mobility of some osteoporotic VCFs
- Include the calculated precision error for all measurements.[154]

He notes elsewhere that the precision error for older, osteoporotic, populations, is in the region of 2% to 5%.[155]

None of the included studies appear to meet McKiernan et al's standards.

Doidge et al note that vertebral height is a surrogate outcome and that, as such, criteria have not been established to determine the clinical importance of any changes.[56]

Angular deformity

Angular deformity may be evaluated in terms of the kyphotic angle, the sagittal index, or measures of sagittal balance.

The kyphotic angle has been defined as the angle defined by the intersection of a line drawn on an x-ray through the posterior-superior and anterior-superior endplate margins of an individual vertebra and a line drawn through the posterior-inferior and anterior-inferior endplate margins of the same vertebra. McKiernan et al reported a precision error of 15.6% in the measurement of the kyphotic angle in elderly osteoporotic patients undergoing PVP or BKP.[156] However, McKiernan notes that most studies which report improvements in the kyphotic angle fail to provide an osteoporosis-appropriate precision error, and states that, in the absence of a measured and reported precision error, the statistical significance of any comparison of kyphotic angles is suspect.[155]

The sagittal index is a measure of kyphotic segmental deformity at the level of a given mobile segment (i.e., one vertebra and one disk), corrected for the normal sagittal contour at the level of the deformed segment.[157] Taking the baseline sagittal contour into account gives the sagittal index a potential advantage as a diagnostic assessment tool.[158] Jiang et al found the sagittal index to have acceptable correlation coefficients for both inter- and intra-observer reliability. However, they also noted that these rates of agreement were lower than those of two other methods of kyphotic angle assessment – the Cobb angle and Gardner angle. It was suggested that the lower level of reliability was observed because the sagittal index includes a smaller area of measurement, thereby maximising differences between measurements.[158]

Further potential sources of measurement variability may include radiograph quality, type of fracture, fracture location, and the position of the radiographic beam relative to the vertebral level in question.[158]

Sagittal spinal balance is a harmonious alignment of the pelvis and spine which allows ease of standing. It can be assessed using lateral radiographs measuring the relationship between the C7 and S1 vertebrae, with greater plumb line deviation of the C7 vertebra representing higher levels of imbalance. Increases in positive sagittal balance have shown strong linear correlations with a variety of health and disability measures.[159]

As Doidge et al note, no criteria have been established to determine the clinical importance of any changes in angular deformity.[56] Moreover, the potential for change in angular deformity is presumably dependent upon the morphology of the vertebral fracture being treated.

- Incidence of new vertebral fractures

Vertebral fractures may be symptomatic or asymptomatic. Symptomatic, or clinical, vertebral fractures cause either sufficient discomfort for the patient to bring them to the attention of a health professional or a measurable loss of height. Their presence can be confirmed by radiographs. However, radiographs can also identify asymptomatic fractures. Some studies undertake routine imaging at follow-up, and thus report radiographically-identified fractures (also termed radiographic or morphometric), which will include both symptomatic and asymptomatic fractures. However, some studies only report clinical fractures.

None of the various different approaches which have been developed to identify radiographic vertebral osteoporotic fractures has been agreed to be the gold standard. The purely qualitative approach, which depends on the visual identification of abnormalities in vertebral shape or height, is a subjective method with poor inter- and intra-rater reliability; however, unlike a purely quantitative method, when performed by an expert it can exclude vertebral abnormalities which are not osteoporotic in origin.[158] More recently, Jiang et al have developed an algorithm-based qualitative approach which aims to facilitate differentiation between osteoporotic fracture and deformity due to other causes.[158] Quantitative methods are more objective and reproducible than qualitative methods, but may identify non-fracture deformities as fractures, whilst failing to recognise mild endplate fractures.[158] However, the number of false positives may be reduced if the definition of incident fracture requires a 20% or greater reduction in anterior, central, or posterior vertebral height.[160] The semiquantitative method developed by Genant et al.[7,161] grades each vertebra according to the visually apparent degree of reduction in vertebral height and area, irrespective of the type of deformity, but also

gives careful attention to changes in vertebral shape, enabling non-fracture deformities to be excluded whilst endplate fractures which are not associated with a 20% reduction in vertebral height can be identified.[162] The semiquantitative method is more objective and reproducible than the qualitative method, but has better specificity and sensitivity than the quantitative method because it reduces the number of false positives while identifying mild deformities which the quantitative method would exclude.[162] However, some researchers claim that the semiquantitative method can be difficult to apply accurately, and that it overestimates fracture prevalence by failing to differentiate adequately between true fractures and non-fracture deformities.[158,163]

Doidge et al note that the relationship between a vertebral fracture and subsequent vertebral fractures is known to be time-dependent – i.e. the risk of subsequent fractures reduces over time. It is also likely that, if PVP or BKP affect the risk of subsequent fracture, they will do so in a time-dependent manner. Consequently, risks estimated in populations which differ in terms of either baseline fracture age or length of follow-up are not directly comparable.[56]

It has been suggested that vertebrae adjacent to those which have been treated by PVP or BKP may be susceptible to subsequent fractures because the treated vertebrae, being stiffer than those which have not been treated, may transmit increased force to adjacent vertebrae.[164] In this context, the most meaningful outcome measure is the proportion of patients who experience at least one clinically important adjacent fracture.[56] Some studies report only adjacent fractures, while others report all incident fractures, whether adjacent or distant; some report the number of patients who suffer fractures, and others only the number of fractures.

There is also potential for confounding of the data in that patients who have had vertebral augmentation and experienced considerable benefit may become more active, at be at a greater risk of fracture than more sedentary patients.

- Progression of treated fracture

Progression of treated fracture is defined in terms of loss of vertebral height. The only study to report this outcome, VERTOS II, defines progression as a loss of vertebral height of 4mm or over, categorising a loss of 4-7 mm as moderate and a loss of over 8mm as severe.[165]

Mortality and adverse effects of treatment

PVP and BKP have been associated with a range of adverse effects. These include:

- Complications related to insertion of a needle, including infection, venous bleeding,[166] and damage to neural or other structures[85]

- Complications related to the leakage of bone cement or the displacement of bone marrow and other material by the cement. Leakage occurs when the cement is not wholly contained by the fractured vertebra but escapes through either the fracture or the track created by the needle. Cement may leak into the paravertebral soft tissues, the spinal canal or neural foramina, the adjacent vertebral disc spaces, or nearby blood vessels.[56] Such leaks may compress the nerve root (causing radiculopathy which may be transient and treatable with NSAIDs or local steroid injections, or, if pain is persistent, may require surgical removal of the cement) or the spinal cord (resulting in myelopathy, and requiring urgent neurosurgical decompression to prevent neurological sequelae including paresis or paralysis).[16,166] They may also result in pulmonary embolism which may be asymptomatic or may result in signs and symptoms such as chest pain, dyspnoea, tachypnoea, cyanosis, coughing, haemoptysis, dizziness and sweating[123] Pulmonary emboli composed of marrow material displaced by the cement appear to be more common than emboli formed of the cement itself. Healthy individuals can tolerate small pulmonary emboli without symptoms, but a large cement leak can lead to pulmonary infarct, and multiple emboli may lead to pulmonary compromise and even death due to respiratory compromise.[166] Sharp and elongated spike-like cement fragments may also perforate blood vessels or the heart.[167]
- Complications relating to balloon rupture in BKP[86]
- Systemic reactions to the bone cement, including hypotension and death.[90] The mechanisms underlying PMMA-induced systemic reactions are not clear: hypotheses include potential toxic, vasodilating, or allergic effects of the cement, as well as possible bone marrow microemboli.[168]
- Complications relating to other aspects of the procedure such as patient positioning and anaesthesia; these include fracture of the rib or sternum in severely osteoporotic patients[169] and systemic infection.[170]

Some of these adverse events are acute (e.g. bleeding at puncture site, local infection, cement leakage, pulmonary embolism); other sequelae are delayed (e.g. adjacent vertebral fracture, cement dislodgment, pyogenic spondylitis). Some are minor, requiring no surgical intervention, whether immediate or delayed. Others are serious, requiring surgical intervention, or resulting in death or significant disability.[16]

It has been noted that the number of cement leaks which are identified is related to the method used to identify them. For example, in a recent retrospective case series of 181 patients with 277 levels treated with PVP, Martin et al[171] found that computed tomography (CT) scanning detected leakage in 149 patients (82%), while procedural dictation and plain radiography

detected leaks in 62 (34%) and 77 patients (50%). The difference in detection rates between each of these methods and CT scans were statistically significant (both, $p<0.01$). CT scanning is therefore viewed as the most sensitive method for detecting cement extrusion, as small pulmonary cement deposits which remain undetected on chest radiographs are readily apparent on CT.[172]

However, the added sensitivity provided by CT may be of limited benefit. For example, in the open-label VERTOS II randomised controlled trial,[173] perivertebral cement leakage was observed in 80% of treated vertebrae, and all leakages remained clinically silent.[167] Although the long-term implications of such small extrusions remain poorly understood, Venmans et al. have argued that some leakage during PVP is difficult to avoid, and that the real issue concerns clinically relevant leakages.[167]

In addition to adverse effects affecting patients, PVP and BKP pose hazards to healthcare professionals. These relate to:

- Exposure to bone cement
- Exposure to radiation.

In addition to the adverse effects of treatment listed above, Medtronic have put forward evidence drawn from large population-based datasets to suggest that vertebral augmentation also has beneficial effects in terms of reduced mortality and morbidity; the reductions in morbidity are not directly captured by the outcomes included in the current review, while the included studies are not large enough to demonstrate significant differences in mortality.

- Key issues

Research has shown that many patients with acute radiographically-diagnosed VCF who receive conservative treatment report a reduction in pain of 50% or more by 6 months.[19] Because of the self-limiting nature of the condition, it is therefore crucial that outcomes are assessed in terms of the difference between the mean changes from baseline in the intervention and control groups in a randomised trial, and not in terms of the mean change from baseline in a single group of patients who have received the intervention.

4.2 Overall aims and objectives of assessment

The aim of this review is to systematically evaluate and appraise the clinical and cost-effectiveness of percutaneous vertebroplasty and percutaneous balloon kyphoplasty in reducing pain and disability in people with osteoporotic vertebral compression fractures in England and Wales.

5. ASSESSMENT OF CLINICAL EFFECTIVENESS

5.1 Methods for reviewing clinical effectiveness and safety

A systematic review was undertaken according to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

5.1.1 Identification of studies

Extensive searches were undertaken with the aim of comprehensive retrieval of studies of clinical and cost-effectiveness relating to the research question.

The search strategy comprised the following main elements:

- searching of electronic databases listed below
- scrutiny of bibliographies of retrieved papers and previous systematic reviews
- contact with experts in the field.

5.1.1.1 Electronic searches

The searches aimed to systematically identify all literature relating to the clinical and cost-effectiveness of percutaneous vertebroplasty and percutaneous balloon kyphoplasty as treatments for osteoporotic compression fractures in men and women of all ages. A comprehensive database of relevant published and unpublished articles was constructed using Reference Manager© software.

Sources searched

The following electronic databases were searched from inception: Medline (Ovid); Medline in Process; CINAHL; EMBASE; EconLit; the Cochrane Library including the Cochrane Database of Systematic Reviews, Cochrane Controlled Trials Register (CCTR), DARE, NHS EED and HTA databases; Science Citation Index (SCI). The searches were conducted in November 2011.

Search terms

The search strategy was developed in collaboration with the information specialist. Search terms included ‘vertebroplasty’, ‘kyphoplasty’, and a broad variety of related clinical terms (e.g. ‘bone void fill*’, ‘vertebral* augmentation*) in order to obtain a wide scope. No bibliographic filters were used. Vocabulary around vertebroplasty/kyphoplasty is limited, therefore few synonyms where available. The searches were simple with an emphasis on sensitivity, utilising both keywords and MeSH/thesaurus terms where available. The Medline

search strategy is provided in Appendix 1. Search strategies for the other databases are available on request.

Search restrictions

Searches were not restricted by language, publication date, or publication type (with exception of removing letters, news, editorials etc). Furthermore, they were not restricted by study design, because studies which did not meet the inclusion criteria for the review of clinical effectiveness might provide relevant information relating to adverse events or be important in identifying further relevant papers and current research.

5.1.1.2 Scrutiny of bibliographies of retrieved papers and previous systematic reviews

The bibliographies of retrieved papers and the MS were scrutinised to identify relevant evidence.

5.1.1.3 Contact with experts in the field

Our clinical expert (DW) in the field was also consulted on whether the search had missed any relevant studies. He believed that all the relevant RCTs had been successfully identified.

5.1.2.1 Inclusion and exclusion criteria

Inclusion criteria

Population

The population comprised people of any age and either gender with painful osteoporotic vertebral compression fractures. Studies which also included participants with non-osteoporotic vertebral fractures of other aetiologies (e.g. fractures associated with trauma, myeloma, or metastatic cancer) were included if data relating to participants with osteoporotic fractures could be extracted separately, or if the proportion of participants with non-osteoporotic fractures was extremely small.

Intervention(s)

Percutaneous vertebroplasty (PVP); percutaneous balloon kyphoplasty (BKP) with or without vertebral body stenting.

Comparator(s)

The interventions themselves, non-invasive management, OPLA, or no treatment.

Outcomes:

The primary outcomes of interest for this appraisal were:

- Health-related quality of life
- Back-specific functional status/mobility
- Pain/analgesic use
- Vertebral body height and angular deformity
- Progression of treated fracture
- Incidence of new vertebral fractures

Secondary outcomes were:

- All-cause mortality
- Symptomatic and asymptomatic leakage of cement (e.g. into adjacent intervertebral discs)
- Periprocedural balloon rupture
- Post-operative complications (including infection)
- Other adverse events
- Resource utilisation
- Cost utility.

Only studies which reported data relating to at least one of the primary outcomes listed above in relation to the population of interest were eligible for inclusion in the review of clinical effectiveness. This criterion was relaxed for consideration of adverse events, to allow the inclusion of studies which reported data relating to any of the secondary outcomes in the population of interest. However, adverse event data were included only if they related to patients with osteoporotic VCFs because of the possibility that patients with fractures of different aetiology (e.g. malignancy) might be susceptible to more, or different, adverse events.

To facilitate comparison, outcomes measured at or before three weeks are grouped together as short-term outcomes, those measured between one and six months as medium-term outcomes, and those measured at 12 months or later as long-term outcomes.

Study design

According to the accepted hierarchy of evidence, the review of clinical effectiveness was limited to randomised controlled trials (RCTs), as they provide the most authoritative form of evidence. It was planned that non-randomised studies would be considered if insufficient data were available from RCTs, but this was not necessary.

In reviews of interventions for which beneficial effects are uncertain or contentious, with some possibility of harm, an accompanying review of adverse events (AEs) can be of substantial importance when deciding whether to use the intervention.[176] It is widely recognised that RCTs do no form a good source of evidence for adverse events: they are generally not powered to reliably detect rare adverse effects, nor is their follow-up period long enough to permit the detection of adverse effects widely separated in time from the original intervention.[177] In addition, their populations are often not wholly typical of the target population: they may be younger and have fewer comorbidities than the general population of patients with the condition of interest.[178] Moreover, RCTs do not always measure all potential side-effects.[179] Hence, it was decided to review the literature on AEs in PVP and BKP to provide additional support for clinical decision making. AEs were addressed using two broad research questions, namely, "what AEs are associated with PVP or BKP in the treatment of osteoporotic VCFs?" and "what is the approximate incidence of AEs associated with PVP or BKP in the treatment of osteoporotic VCFs?" Although this broad approach risked an overload of heterogeneous data which could not be easily pooled, it had a twofold advantage: first, it could identify new or previously unrecognised complications, and second, it could provide a more comprehensive overview of potential complications.[176]

Two types of evidence were included in the review of safety:

- Large observational studies ($\geq$200 patients), which would allow exploration of the range and incidence of adverse events associated with PVP and BKP. The decision to include only large observational studies was based on a desire to exclude small case series which might display particularly high adverse event rates associated with limited experience of the relevant techniques on the part of the clinician or institution. The decision to set the threshold for inclusion at 200 patients was taken *a priori*.
- Individual case reports were used to supplement the RCTs and large observational studies to provide as full a picture as possible of the range of adverse events associated with PVP and BKP. They were therefore used as a source of evidence relating only to adverse events which were not reported in the RCTs or large observational studies. By their nature, individual case reports cannot provide any indication of the incidence of such adverse events.

As with the review of clinical effectiveness, studies which included participants with non-osteoporotic vertebral fractures of other aetiologies (e.g. fractures associated with trauma, myeloma, or metastatic cancer) as well as those with osteoporotic VCFs were excluded unless data relating to participants with osteoporotic fractures could be extracted separately. This was because there is some evidence that the type and incidence of AEs may differ in vertebral fractures of non-osteoporotic origin (e.g. metastatic, traumatic).[180,181]

Exclusion criteria

Systematic reviews were excluded from the review of clinical effectiveness and safety, but were retained for discussion and identification of additional relevant primary research studies. Studies which were considered methodologically unsound were excluded from the review, as were the following publication types:

- Animal models
- Preclinical and biological studies
- Narrative reviews, editorials, and opinions
- Non-randomised studies (except for adverse effects)
- Studies published as meeting abstracts only, which reported insufficient methodological details to allow critical appraisal of study quality.

In addition, potentially relevant publications were excluded if they had been superseded by later publications and did not contain any additional useful data: this applied to several conference abstracts.

Study selection

Retrieved studies were selected for inclusion through a two-stage process according to the above inclusion/exclusion criteria. The references identified by the literature searches were assessed for relevance first by title/abstract, and then by full text, excluding at each step studies which did not satisfy those criteria; abstract-only publications were retained for full-text review. One reviewer examined titles and abstracts for inclusion, and screening was checked by a second reviewer on ten per cent of citations. For studies of clinical effectiveness, the kappa coefficient (range 0 to 1) calculated to measure inter-rater reliability was excellent, at 1.0, indicating no discrepancies.

Data extraction strategy

Data were extracted independently by two reviewers using a standardised data extraction form (see Appendix 2); discrepancies were resolved by discussion and did not require input

from a third reviewer. Where multiple publications relating to the same study were identified, data were extracted and reported as a single study.
Data obtained from the submissions made by the manufacturers have been appraised and commented on where deemed relevant.

Critical appraisal strategy

The methodological quality of each study which met the inclusion criteria for the review of clinical effectiveness was assessed independently by two reviewers, and any discrepancies were resolved by discussion. Where a study was reported in more than one publication, its quality was assessed on the basis of the combined data from all relevant publications.

It was stated in the protocol that quality would be assessed according to criteria based on those proposed by Ploeg et al. for the assessment of studies of percutaneous vertebroplasty[87] (see Appendix 3). These criteria were initially adopted because they could be applied to both randomised and non-randomised studies. However, in the event, because sufficient RCTs were identified, it was not necessary to include non-randomised studies, and it was found that the criteria proposed by Ploeg et al. did not discriminate sufficiently between the included RCTs. A new set of criteria was therefore developed: this was based on the criteria proposed by CRD and the Cochrane Collaboration for assessing the risk of bias in randomised trials, but also incorporated criteria proposed by Ploeg et al.[87] and Furlan et al.[133] which had particular relevance to the interventions under review. These criteria relate to internal validity, and also to external validity, and precision (for details, see Appendix 4). The criterion relating to the blinding of care providers was excluded as such blinding was not possible given the nature of the interventions under review.

The revised quality assessment tool included some questions which led to subsidiary questions to which the answer could be “not applicable”. These subsidiary questions have not been included in the risk of bias tables presented in section 5.2.2.

Methods of data synthesis

Due to the potential impact of baseline imbalances in the degree of pain and disability reported by patients with osteoporotic VCFs, it is crucial that outcomes which are reported as continuous data (e.g. pain, disability, and health-related quality of life) are assessed in terms of the difference between the mean changes from baseline in the intervention and control groups, and not in terms of the differences between mean scores at any given point in time. Where the original study investigators presented relevant measures of effect in terms of mean changes from baseline, these have been included in the data tables not least because in some

cases they also adjusted for stratification variables (e.g. treatment centre). Where adjusted data were not reported, mean between-group differences in change from baseline for continuous outcomes were calculated adjusting for the variance of the within treatment change from baseline, where this was made possible by the data. This method generated confidence intervals but not p values. For dichotomous outcomes, relative risks, with confidence intervals and p values, were calculated using the Cochrane Collaboration Review Manager© Software (version 5.1) if such data were not reported by the study investigators.[182]

Studies which met the review's entry criteria were eligible for inclusion in meta-analyses to estimate summary measures of effect if such meta-analysis was appropriate (i.e. if the study populations, intervention, and outcomes were comparable). Meta-analysis was carried out using random effects models, using Review Manager© Software (version 5.1).[182] Heterogeneity in the meta-analyses was explored through consideration of the study populations, methods, and interventions, by visualisation of results and, in statistical terms, by the χ^2 test for homogeneity and the I^2 statistic. However, such meta-analysis was limited to dichotomous outcomes.

The review team did not undertake meta-analyses of data relating to continuous or quasi-continuous outcomes. Such meta-analysis was considered inappropriate because of the existence of a published meta-analysis by Staples et al[109] of data from the only two double-blind studies of vertebral augmentation (Buchbinder[101] and INVEST[102]). As this meta-analysis used individual patient data (IPD), it was of a higher quality than could be achieved using published data. There was considered to be too much heterogeneity to justify combining data from all the studies of PVP in a meta-analysis together with published data from the Buchbinder and INVEST studies.

5.2 Results

5.2.1 Quantity and quality of research available

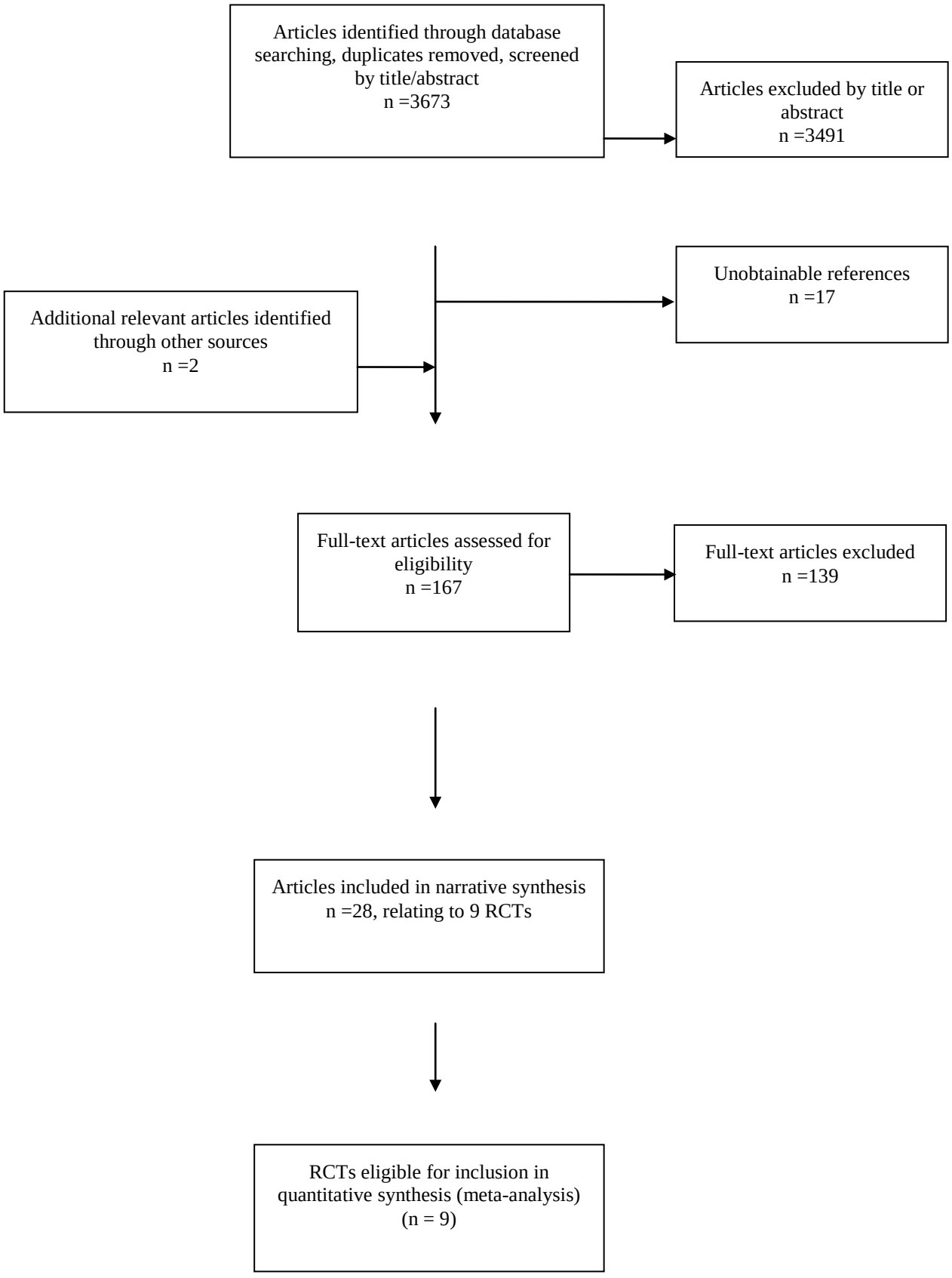
5.2.1.1 Number of studies identified

The electronic literature searches identified 3674 potentially relevant citations. For the review of clinical effectiveness, 3491 of these citations were excluded at the title or abstract stage, leaving 165 which were obtained for examination of the full text.

139 citations were excluded at the full text stage. A further 17 could not be obtained within the study timescale: as almost all of these appear to have been conference abstracts (for details, see Appendix 5), it seems unlikely that potentially valuable information has been

missed as a result of their exclusion. Two additional papers were identified from other sources: these related to the studies by Blasco et al[146]and Rousing et al.[139] Other publications from these studies[183-185] had been identified by the electronic searches. Thus, 28 articles relating to a total of nine RCTs were identified and included in the review of clinical effectiveness (see Figure 1).

Figure 1: Clinical Effectiveness: Summary of Study Selection and Exclusion



In their systematic review of vertebroplasty for painful osteoporotic VCFs, Muijs et al.[186] identified a further two RCTs which were not identified by the electronic literature searches: a small study by Do et al,[187] and a pilot study by Kallmes et al,[188] both of which have only been published as conference abstracts. Neither of these studies met the inclusion criteria for the current review. Do et al randomised 31 patients to either PVP or continued medical therapy, but the latter group received PVP six weeks after the PVP group, and only before-and-after data are presented for each group: thus, no comparison was drawn between treated and untreated patients, or indeed between patients in whom PVP was performed sooner or later. Furthermore, baseline data from the control group appear to have been collected immediately before PVP (i.e. 6 weeks later than in the original intervention group), but this is not clear.[187] The study by Kallmes et al was a pilot study intended to demonstrate the feasibility of enrolling patients into a trial of percutaneous vertebroplasty against OPLA. Only five patients were enrolled; although pain was used as an outcome, it was not quantified, and patients were only said to have gained "minimal pain relief" or "complete pain relief".[188]

5.2.1.2 Number and type of studies included

A total of 9 RCTs met the review inclusion criteria. These compared:

- percutaneous vertebroplasty with an OPLA (Buchbinder 2009, INVEST)
- percutaneous vertebroplasty with optimum pain medication (Farrokhi 2011, VERTOS, VERTOS II)
- percutaneous vertebroplasty with conservative treatment (Blasco 2012, Rousing 2009)
- percutaneous balloon kyphoplasty with non-surgical management (FREE)
- percutaneous balloon kyphoplasty with percutaneous vertebroplasty (Liu 2010).

Details of the techniques for vertebral augmentation in these trials is included in the Appendix 8, table 100. For simplicity, we have used the term optimal pain management (OPM) as a term encompassing conservative treatment and non-surgical management.

The principal source/sources for each study are listed in Table 2; a full list of publications relating to each study is included in Appendix 6.

Table 2: Principal sources for each trial

Trial name and identifier	Primary Report(s)
Blasco 2012 (NCT00994032)	Blasco et al 2012[146]
Buchbinder 2009 (ACTRN012605000079640)	Buchbinder et al 2009[101]
Farrokhi 2011 (IRCT138804252193N1)	Farrokhi et al 2011[147]
FREE (NCT00211211)	Wardlaw et al 2009[151] Boonen et al 2011[196]
INVEST (NCT00068822)	Kallmes et al 2009[102]
Liu 2010	Liu et al 2010[148]
Rousing 2009	Rousing et al 2009[139] Rousing et al 2010[185]
VERTOS	Voormolen et al 2007[153]
VERTOS II (NCT00232466)	Klazen et al 2010[17]

It should be noted that the FREE study has been included even though, strictly, it does not meet the inclusion criteria because it included four patients with multiple myeloma. However, as these patients formed only 1% of the study population and were evenly distributed between treatment groups, it seemed unreasonable to exclude the study from the review in the absence of any other RCT comparing BKP with non-invasive management.

5.2.1.3 Number and type of studies excluded

As may be seen from section 6.2.1, a substantial number of the citations identified by the electronic searches related to studies which were excluded as part of the sifting process because they did not meet the inclusion criteria. Details are therefore given only of those citations which were excluded after a full reading, and then only if they were excluded for a reason other than a simple failure to meet the inclusion and exclusion criteria. Such citations are listed in Appendix 7 together with the reasons for their exclusion.

5.2.1.4 Ongoing or unpublished trials

Six relevant trials were identified from the NCT website (http://clinicaltrials.gov/ct2/): NCT00203554 compared PVP with conservative treatment, VERTOS IV compared PVP with OPLA, and the remaining four compared BKP with PVP and, in one case (OSTEO-6), also with conventional treatment (for details, see Table 3). One of these trials (NCT00203554) has been completed, but the results have not been released because they are currently submitted to

a journal (Leif Sørensen pers. comm. 28.2.12). The NCT website previously said that the KAVIAR Trial (NCT00323609) was ongoing but not recruiting participants, and had an estimated primary completion date of August 2011 (access date 21.09.2011); however, it now says that it has been terminated[189] and the Medtronic submission states that the results of a partial analysis are expected in July 2012.[34]

Two additional ongoing studies were identified which do not appear to be included in the NCT website:

- a double-blind study by Longo et al,[190] identified by the Embase search, comparing PVP with conservative treatment
- STU-SPI-S-06-134-01, identified by the Synthes submission,[32] comparing percutaneous stentoplasty with BKP.

Despite the existence of VERTOS II and VERTOS IV, no VERTOS III could be identified.

Table 3: Ongoing or unpublished randomised trials of percutaneous vertebroplasty or percutaneous balloon kyphoplasty in patients with painful osteoporotic vertebral compression fractures

Study	Intervention	Comparator/s	Sponsor	Status
NCT00203554[191]	Percutaneous vertebroplasty	Conservative treatment of pain	University of Aarhus	Completed but unpublished
CEEP (NCT00279877)[192]	Percutaneous balloon kyphoplasty	Percutaneous vertebroplasty	Mayo Clinic	Said to be ongoing, with an estimated primary completion date of Aug 2011. First results expected March 2012[34] but not posted on NCT website as at 15.5.12
KAVIAR Trial (NCT00323609)[189]	Percutaneous balloon kyphoplasty	Percutaneous vertebroplasty	Medtronic Spine LLC	Terminated but unpublished
OSTEO-6 (NCT 00749060)[193]	Percutaneous balloon kyphoplasty	Percutaneous vertebroplasty Conventional treatment	Assistance Publique – Hôpitaux de Paris	Estimated completion date Dec 2012
OSTEO+6 (NCT00749086)[194]	Percutaneous balloon kyphoplasty	Percutaneous vertebroplasty	Assistance Publique – Hôpitaux de Paris	Estimated completion date Dec 2012
VERTOS IV (NCT01200277)[195]	Percutaneous vertebroplasty	Sham procedure	St. Elisabeth Hospital, Tilburg, Netherlands	Recruiting; data collection ongoing until Jan 2013
Longo et al[190]	Percutaneous vertebroplasty	Conservative treatment (defined as 3 weeks' bed rest wearing a rigid hyperextension suspension brace followed by 2-3 months wearing a Cheneau brace)	None reported	Not clear
STU-SPI-S-06-134-01[32]	Percutaneous stentoplasty	Percutaneous balloon kyphoplasty	Synthes GmbH	Results expected to be available by the end of 2013

5.2.2 Study characteristics

A summary of the characteristics of the included studies is provided in Table 4. Tables providing details of the technical characteristics of the PVP and BKP procedures, and the reporting of clinical outcomes, may be found in Appendix 8. Baseline demographic data are presented in Table 5. Only three studies (Farrokhi,[147] VERTOS,[153] VERTOS II[17]) defined osteoporosis in terms of bone mineral density (BMD); the remainder appeared to assume the presence of osteoporosis from the presence of VCF in the absence of any other known fracture aetiology.

Table 4: Characteristics of included studies

Study	Country	Recruitment dates	Total numbers randomised	Length of follow-up	Lost to follow-up	Intervention	Comparator	Crossover to intervention permitted	Source of funding
Blasco 2012[146]	Spain	April 2006 to January 2010	125	12 months	24%	Percutaneous vertebroplasty plus nasal calcitonin for 1 month	Conservative treatment (standardised analgesia; nasal calcitonin for 1 month; rescue therapy with intrathecal infusion if necessary)	Yes	Fundació La Marató de TV3; Spanish Society of Medical Radiology; Catalan Society of Rheumatology
Buchbinder 2009[101]	Australia	April 2004 to October 2008	78	2 years. However, data only available for 6 months	9%	Percutaneous vertebroplasty	Sham procedure without local anaesthesia	No	National Health and Medical Research Council of Australia (284354); Cabrini Education and Research Institute; Cook Australia
Farrokhi [147]	Iran	Sept 2004 to Jan 2006	82	3 years	7%	Percutaneous vertebroplasty	Optimal medical therapy (suggested baseline analgesia 250 mg paracetamol + codeine and 400 mg ibuprofen, both twice daily; also 1000 mg calcium, 400 IU vitamin D, and 200 IU calcitonin daily; and 70	Yes after 1 month	Shiraz University of Medical Sciences; Apadana Tajhizgostar Co

Study	Country	Recruitment dates	Total numbers randomised	Length of follow-up	Lost to follow-up	Intervention	Comparator	Crossover to intervention permitted	Source of funding
							mg oral alendronate once a week)		
FREE[151]	Multinational (Austria, Belgium, France, Germany, Italy, The Netherlands, Sweden, UK)	Feb 2003 to Dec 2005	300	24 months[196]	23%	Balloon kyphoplasty	Non-surgical management according to local practice (also provided to intervention group)	No; patients who wished to cross over were withdrawn from the study	Medtronic Spine LLC
INVEST[102]	Multinational (USA, UK, Australia)	June 2004 to August 2008	131	Intended to be 12 months.[152] However, data only available for 3 months	5%	Percutaneous vertebroplasty	Sham procedure with local anaesthesia	Yes after 1 month	National Institute of Arthritis and Musculoskeletal and Skin Diseases
Liu 2010[148]	Taiwan	NR	100	Minimum of 6 months	NR	Balloon kyphoplasty	Percutaneous vertebroplasty	No	Chung-Shan Medical University Hospital
Rousing 2009[139]	Denmark	Jan 2001 to Jan 2008	50	12 months	10%	Percutaneous vertebroplasty	Conservative treatment (pain medication and physiotherapy until discharge, as in intervention group; in addition, bracing also offered)	No	Foundation and Danish government funding
VERTOS[153]	Belgium and The	July 2003-June	46	Intended to be	26%	Percutaneous	Optimum pain medication	Yes at 2	NR

Study	Country	Recruitment dates	Total numbers randomised	Length of follow-up	Lost to follow-up	Intervention	Comparator	Crossover to intervention permitted	Source of funding
	Netherlands	2005		1 year. However, the study was stopped prematurely at 2 weeks because most control patients asked to cross over to PVP		vertebroplasty	(in ascending order of anaesthesia: paracetamol, NSAIDs, or opiate derivatives, according to individual need)	weeks	
VERTOS II)[17]	Belgium and The Netherlands	1.10.05 - 30.6.08	202	12 months	19%	Percutaneous vertebroplasty	Optimum pain medication (also offered to intervention group); in addition, physiotherapy or bracing also offered	Yes, apparently after 1 week	ZonMw (Dutch organisation for health care research and innovation of care),project number 945-06-351; COOK Medical

Table 5: Baseline demographic data

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
Blasco 2012[146]	71.33+9.95/ 75.27±8.53	47/64 (73%)/ 50/61 (82%)	Lumbar: -2.48±1.77/ -2.80±1.32 Femoral neck: -2.14±0.97/ -2.24±0.87	Duration of pain; bone marrow oedema seen on MRI or activity on bone scan	<12 months	Mean duration of back pain (days): 140.3±96.09/ 143.1±130.33 *In weeks: 20.0±13.7/ 20.4±18.6*	≥4 7.21±0.33/ 6.31±0.35	*47*/64 (*73%*)/ *31*/61 (*51%*)	Unspecified osteoporosis: 100%
Buchbinder 2009[101,150]	74.2±14.0/ 78.9±9.5	31/38 (82%)/ 31/40 (78%)	NR	Duration of pain; bone marrow oedema, a fracture line, or both seen on MRI) (if	<12 months	Median duration of back pain (weeks): 9.0 (3.8-13.0)/ 9.5 (3.0-17.0)	No minimum requirement specified 7.4±2.1/ 7.1±2.3	30/38 (79%)/ 34/40 (85%)	Unspecified osteoporosis: 100% Baseline corticosteroid use: 37%

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
				MRI not feasible, fracture identified by CT scan and increased uptake compatible with recent vertebral fracture assessed by bone scan)					
Farrokhi [147]	72 (range 59-90)/ 74 (range 55-87)	30/40 (75%)/ 30/42 (71%)	T-score <-2.5	Duration of pain; vacuum phenomenon or bone	4 weeks-12 months	Median duration of back pain (weeks): 27 (4-50)/	No minimum requirement specified	30/40 (75%)/ 30/42 (71%)	Primary osteoporosis: 100%

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
				marrow oedema seen on MRI		30 (6-54)	8.4±1.6/ 7.2±1.7		
FREE [151]	72.2 (9.3)/ 74.1 (9.4)	115/149 (77%)/ 117/151 (77%)	NR	Duration of pain; bone marrow oedema seen on MRI	<3 months	Mean duration of back pain (weeks): 5.6 (4.4)/ 6.4 (5.2)	≥4 NR	103/140 (74%)/ 99/146 (68%)	Primary osteoporosis: 96% Secondary osteoporosis: 3% Multiple myeloma/ metastatic: 1% Baseline corticosteroid use: 17%

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
INVEST[102,152]	73.4±9.4/ 74.3±9.6	53/68 (78%)/ 46/63 (73%)	NR	Duration of pain; if unclear, bone marrow oedema on MRI or increased vertebral-body uptake on bone scanning	<12 months	Mean duration of back pain (weeks): 16 (IQR 10-36)/ 20 (IQR 8-38)	≥3 6.9±2.0/ 7.2±1.8	38/68 (56%)/ 40/63 (63%)	Unspecified osteoporosis or osteopaenia: 100%
Liu 2010[148]	72.3±7.6/ 74.3±6.4	39/50 (78%)/ 38/50 (76%)	NR	NR	NR. PV or BKP said to have been performed within 43 days of injury	Mean duration of back pain (days): 17.0±7.7/ 15.8±6.7 *In weeks:*	No minimum requirement specified 8.0±0.8/ 7.9±0.7	NR	Unspecified osteoporosis: 100%

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
						2.4 (1.1)/ *2.3 (1.0)*			
Rousing 2009[139]	80 (range 65-96)/ 80 (range 71-93)	19/25 (76%)/ 21/24 (88%)	NR	Duration of pain; if patients had >1 fracture, fractures were accepted as new if they showed bone marrow oedema on MRI or increased bone turn-over on bone scan	<8 weeks	Mean fracture age (days): 8.4 (95% CI 3.7-13.0)/ 6.7 (95% CI 2.1-11.4) *In weeks:* *1.2 (0.5-1.9)/* *1.0 (0.3-1.6)*	No minimum requirement specified 7.5 (95% CI 6.6-8.4)/ 8.8 (95% CI 8.2-9.3)	NR	Unspecified osteoporosis: 100%
VERTOS[153]	72 (59-84)/	14/18 (78%)/	T-score <-2.0	Duration of	6 weeks-6	Mean duration	No minimum	6/18 (66%)/	Unspecified

Study	Mean age in years (intervention/ comparator)	Number female (intervention/ comparator)	BMD	Method used to assess fracture age	Acceptable duration of fracture pain	Time from estimated fracture onset to intervention (intervention/ comparator)	Minimum pain score on 0-10 scale required for inclusion Mean baseline pain score (0-10 scale) (intervention/ comparator)	Baseline opioid analgesia use	Fracture aetiology
	74 (55-88)	14/16 (88%)		pain; bone marrow oedema seen on MRI	months	of pain (days) (range): 85 (47-138)/ 76 (46-141) *In weeks:* *12.1 (6.7-19.7)/* *10.9 (6.6-20.1)*	requirement specified 7.1 (5-9)/ 7.6 (5-10)	5/16 (31%)	(implicitly primary) osteoporosis or osteopaenia: 100%
VERTOS II)[17]	75.2 (9.8)/ 75.4 (8.4)	70/101 (69%)/ 70/101 (69%)	T-score ≤-1	Duration of pain; bone marrow oedema seen on MRI	≤6 weeks	Mean duration of pain (days): 29.3±17.1/ 26.8±16.0 *In weeks:* *4.2±2.4/* *3.8±2.3*	≥5 7.8±1.5/ 7.5±1.6	50/95 (53%)/ 44/92 (48%)	Unspecified osteoporosis or osteopaenia: 100%
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers									

Study quality

Internal validity

The included studies varied in terms of internal validity (see Figure 2). The potential sources of bias are discussed in turn below.

Figure 2: Risk of bias summary: review authors' judgements about each risk of bias item for each included study (+ = low risk; - = high risk; ? = unclear risk)

Risk of selection bias

All the included studies were described as randomised. However, as Blasco et al stated only that "randomisation was done with a previously defined randomised computer list",[146] there was lack of clarity about the method of both assignment to treatment groups and concealment of allocation. Liu et al provided no information regarding concealment of allocation.[148]

In multi-centre trials of interventional procedures, stratification of randomisation by treatment centre is important to avoid potential imbalances associated with differences in techniques and skills. Two of the five multi-centre studies (Buchbinder, INVEST) specified that randomisation was stratified by treatment centre; in a third (FREE), although randomisation was stratified, treatment centre was not one of the variables. In VERTOS and VERTOS II, randomisation is not said to have been stratified. Of the remaining four studies, that by Blasco et al was definitely a single-centre study, and those by Liu and Rousing were probably also single-centre. In the study by Farrokhi et al, although two hospitals appear to have been involved, the same surgeon seems to have undertaken the procedure in both.[197] Thus, the risk of bias due to differences in techniques and skills appears to be higher in FREE, VERTOS, and VERTOS II than in the remaining studies.

Risk of performance bias

Few studies appear to be at risk of bias associated with co-interventions as they offered the same oral analgesics to both the treatment and control groups. In the study by Blasco et al, the methods section suggests that rescue therapy by intrathecal infusion (25 μg fentanyl and 1.5 mg bupivacaine) was only offered to patients in the control group in whom drug therapy had proved ineffective (VAS >7) or intolerable. If there was then no improvement in pain, conservative treatment was deemed to have failed and the patient was considered for PVP. However, it is clear from the results section that this rescue therapy was also made available to patients who had received PVP.[146] The risk of performance bias was unclear in the study by Rousing et al: this did not explicitly define conservative treatment or detail the steps taken to avoid co-interventions, but appeared to offer brace treatment to the control group but not the PVP group.[139]

Some of the studies were at risk of bias because of crossover – i.e. patients who were randomised to one intervention receiving the other intervention. Crossover is generally more common in unblinded than in blinded studies because, in unblinded studies, the

patients in the control group are aware that they have not received the experimental intervention. Three of the included studies appear to be free of crossover: it was not allowed in the studies by Buchbinder and Rousing, and was presumably impossible in Liu et al's comparison of PVP with BKP. In the remaining six studies, patients randomised to one treatment might choose, after a minimum period of time, to receive the other treatment; in the FREE study, they were then considered to have withdrawn from the study, but were included in the intention to treat analysis at one year.[151] In the blinded INVEST study, patients were informed at recruitment that they would be allowed to cross over to the other procedure a month or more after the intervention if adequate pain relief was not achieved; specific numerical pain thresholds were not used to determine the adequacy of pain relief and therefore eligibility to cross over.[102] Although crossover was permitted, blinding was maintained for the full year.[152] In the unblinded studies, crossover was reported to be in one direction only, from conservative treatment to the intervention; however, in the Blasco, FREE, and VERTOS II studies, some patients allocated to PVP or BKP did not receive the intervention for various reasons, and the investigators did not appear to adjust for non-treatment. The blinded INVEST study reported crossover from PVP to the OPLA as well as from the sham procedure to PVP: by 3 months, 8/68 (12%) of patients allocated to PVP had crossed over, compared with 27/63 (43%) allocated to OPLA[102] (see Table 6).

Table 6: Crossover from allocated treatment groups

Study	Allocated to PVP but did not receive surgery, or crossed over to control	Allocated to BKP but did not receive surgery	Allocated to control intervention but crossed over to receive PVP or BKP	Time point to which data uncontaminated by crossover
Blasco 2012[146]	7/64 (11%) (5 refused PVP, 2 improved spontaneously)	N/A	7/61 (11%)	Data contaminated from baseline by patients in PVP group who refused intervention; not clear when patients crossed over from the control group
Buchbinder 2009[101]	0	N/A	0	Data appear uncontaminated to end of follow-up
Farrokhi [147]	0	N/A	10/42 (26%)	Data uncontaminated to 1 month
FREE[151]	N/A	10/149 (7%)	14/151 (9%)	Data contaminated from baseline by patients in BKP group who refused intervention; some patients crossed over from control group at <1 month
INVEST[102]	8/68 (12%) (crossed over to control intervention)	N/A	27/63 (43%)	Not clear: crossover not permitted before 1 month but 1 patient in the VPV group and 2 in the control group underwent the crossover intervention at <1 month
Liu 2010[148]	0	0	N/A	Data appear uncontaminated to end of follow-up
Rousing 2009[139]	0	N/A	0	Data appear uncontaminated to end of follow-up
VERTOS[153]	0	N/A	14/16 (88%)	Data appear uncontaminated up to 2 weeks
VERTOS II[17]	8/101 (8%) (2 withdrew consent, 3 refused PVP, 3 improved spontaneously)	N/A	15/101 (15%)	Data contaminated from baseline

Several studies which were theoretically at risk of bias due to crossover attempted to avoid such bias by only reporting results up to the point at which crossover was permitted. Thus, the INVEST study, which permitted crossover after 1 month, only reported comparative results up to that point; 2/42 patients (5%) appear to have crossed over from OPLA to PVP before 1 month.[102] The VERTOS study, which permitted crossover at 2 weeks, was stopped prematurely at that point because 88% of the control group (14/16) requested PVP.[153] In the study by Blasco et al, 7/61 patients allocated to conservative therapy (11%) underwent PVP,[146] but the date at which this occurred was not reported. Farrokhi et al permitted crossover after 1 month; 10 of the 42 control patients (24%) had undergone PVP by 1 year, and a further 10 apparently underwent PVP after 1 year.[147] In the FREE study, 14/151 patients in the control group (9%) underwent BKP, nine of them (6%) before one month.[151] In VERTOS II, 5/101 patients withdrew from the control group before treatment because they wanted PVP; because they withdrew consent, the vertebroplasty procedure could not be documented and analysed in those patients. A further 10 control patients (10%) requested PVP at various points during the study.[17] Thus, the studies with the highest rates of crossover (INVEST, VERTOS) only report comparative results for the period preceding any substantial crossover, and the 1-month data from the study by Farrokhi et al are also unaffected. However, there is some potential for bias due to crossover in the study by Blasco et al and the FREE and VERTOS II studies.

Only two studies (Buchbinder, INVEST) sought to avoid bias by blinding patients to their treatment allocation; both used an OPLA to do so. In the INVEST study, the success of blinding was evaluated: at 14 days, 51% of patients in the PVP group and 63% in the control group correctly guessed their allocation, but in both cases their degree of confidence in the accuracy of their guess was only moderate.[102] Buchbinder et al planned to evaluate the success of blinding at the end of the study,[150] but do not appear to have reported the results. In the study by Liu et al, patient blinding was presumably feasible, since the patients received PVP or BKP under general anaesthesia, but, as patient blinding was not mentioned, it seems unlikely that it was done, particularly as it was stated that some outcomes were assessed by blinded assessors.[148] The remaining studies made no attempt to blind patients to their treatment allocation. Consequently, as many of the outcomes were patient-reported and subjective in nature, the risk of bias associated with lack of blinding in these studies is substantial.

Risk of detection bias

Because of the radio-opaque nature of the cement used for PVP and BKP, it is impossible to blind the assessors of radiographic outcomes (vertebral body height, kyphotic angle, and incident fracture) to treatment allocation. However, there is no reason why blinded assessors should not have been used to collect data relating to other outcomes, yet only four studies (Buchbinder,[101] Farrokhi,[147] INVEST,[102] Liu[148]) stated that they used blinded outcome assessors for at least some outcomes. Three studies (Blasco,[146] FREE,[151] VERTOS II[165]) stated that the radiopacity of the bone cement made it impossible to blind outcome assessors, and appeared to make no attempt to use blinded assessors for non-radiological outcomes. There appeared to be no blinding of outcome assessors in the Rousing and VERTOS studies. As many of the non-radiological outcomes were subjective patient-reported outcomes, which may be modified by contact with non-blinded outcome assessors, such data are at risk of bias in all except the Buchbinder and INVEST studies, and possibly also that by Liu et al.

In the majority of studies, outcomes appeared to be assessed at comparable times in both groups. However, in the study by Liu et al, the timing of the assessment of some outcomes was not clear.[148] In VERTOS II, the timing of assessment of both baseline characteristics and subsequent outcomes appears not to be comparable because baseline was said to be the day of randomisation for the control group but the day of PVP for the intervention group; moreover, PVP was said to have been performed a mean of 9.4 (SD 8.1) days after randomisation.[17] This has two major consequences:

- although the inclusion criteria stipulate that participants should have had back pain for no more than 6 weeks, many patients in the PVP group would have undergone the intervention more than 6 weeks after pain onset, and thus would have subacute rather acute VCFs, whereas all of the control group would have had acute fractures at baseline. Consequently, Doidge et al suggest that between-group differences in baseline variables such as the EQ-5D, which Klazen et al ascribe to chance, may in fact be due to differences in fracture acuity[56]
- in the control arm of VERTOS II, the mean pain score fell by 1.9 points (25%) during the first week following randomisation.[17] A comparable reduction might presumably be expected in the PVP group between randomisation and the assessment of baseline characteristics on the day of the intervention. However, as the “baseline” pain score was 7.8 in the PVP group compared with 7.54 in the control group, if such a reduction occurred, it implies both a substantial

disparity between groups at randomisation and an implausibly high mean pain score in the PVP group at that point in time.

There is thus considerable lack of clarity in relation to the timing of assessments in VERTOS II, and clarification has been sought, but not received, from the study first author.

Risk of attrition bias

Five studies (Buchbinder, Farrokhi, INVEST, Rousing, VERTOS II) followed up at least 80% of participants. Liu et al made no reference to attrition: this may be because there was none, but this is not specified. In the Blasco and FREE studies, only 76% and 78% respectively completed follow-up at 12 months; in the FREE study, there was a marked disparity between treatment groups, with 83% in the BKP group completing follow-up at 12 months, compared with 74% in the control group,[151] whereas in the study by Blasco et al follow-up was higher in the control group (79% vs. 73%).[146] In the VERTOS study, only 74% overall completed 2 weeks follow-up. The data are poorly presented, making a full comparison of drop-out rates in the two treatment groups impossible: 6 patients are said to have withdrawn from the control group because they wanted PVP, and 2 patients from the PVP group because they wanted the control intervention, but details are not given of the treatment allocation of the 4 patients who refused to complete questionnaires at 2 weeks, nor are their outcomes reported at 1 day.[153]

All studies except Liu et al gave information relating to the reasons for withdrawal. While the absence of such information in Liu's study may further suggest that there were no withdrawals, this was not explicitly stated.

With the exception of VERTOS, all studies specified how many patients were randomised to each group. Moreover, all but Liu stated clearly how many were included in the final analysis, and it is possible that Liu did not provide this information because there were no withdrawals. The majority of studies, including all the studies which reported crossover, reported using intention-to-treat analyses.

Risk of reporting bias

Only three studies (Blasco, Farrokhi, VERTOS II) appeared to be free of selective reporting. The Buchbinder, FREE and INVEST studies reported most but not all of the clinical outcomes specified in the study protocol. It is understood that the longer-term outcomes from the Buchbinder study (presumably including the incidence of new

vertebral fractures at 12 and 24 months), although not yet reported, are to be published; however, this does not explain the failure to report the results of the timed 'Up and Go' test and the patients' perceptions of fatigue and overall health. The FREE study did not report functional outcome and the results of objective functionality tests or data relating to VBH (for details, see Appendix 8, Table 101); the last of those omissions is the most surprising, given that one of the particular merits of BKP is said to be its effect on VBH. The INVEST study did not report adjacent fractures at12 months or implant-related inflammation. No study protocol could be found for the Liu, Rousing, or VERTOS studies, and therefore the risk of reporting bias in these studies is not clear.

Risk of other bias

Studies may be at risk of bias if their intervention and control groups differ in baseline factors which are strongly related to outcome measures.[198] In four studies (Buchbinder, Farrokhi, FREE, INVEST), the treatment groups appeared to be comparable at baseline. Liu et al reported so few baseline characteristics that it was difficult to assess comparability. The remaining studies appeared to be potentially at risk of bias because of differences between groups at baseline. In the study by Blasco et al, mean baseline pain, opioid use, and QUALEFFO scores were lower in the control group than in the PVP group. The investigators, who subdivided opioid use into major and minor opiate derivatives, stated that none of the differences were statistically significant;[146] however, if the data are aggregated, total opioid use is significantly higher in the PVP group than in the control group (RR 1.08, 95% CI 1.08-1.93, p=0.01). Rousing et al noted that the mean baseline pain score was significantly lower in the PVP group than in the control group (7.5 vs. 8.8, p=0.02); no information was presented on baseline analgesic use. Rousing also reported a statistically significant difference in EQ-5D scores, favouring PVP; there were also noticeable between-group differences, favouring PVP, in all domains of the DPQ. Despite these differences, results were not reported as changes from baseline, and statistical significance was attributed to the unadjusted data.[139] In VERTOS, the number of treated fractures was significantly higher in the PVP group (p=0.04), and there were also significantly more wedge fractures, and fewer biconcave fractures, in the PVP group than in the control group (p=0.02);[153] the potential impact of these differences on the success of PVP is not clear. Finally, in VERTOS II, there were said to be significant differences between groups in EQ-5D, QUALEFFO, and RDQ scores;[17] in each case, the status of the PVP group was worse than that of the control group. However, adjusted results were reported.

External validity (generalisability) and precision

External validity

Most of the included studies specified their eligibility criteria, thus enabling assessment of the nature of their patient populations. However, many reported that a substantial proportion of patients declined to participate, and only Farrokhi et al specifically stated that the rate of refusal to participate was low[147] (see Table 7). It is not always clear whether patients refused to participate before or after they were found to meet the study inclusion criteria. Consequently, the figures included in Table 7 are presented as percentages of the total number of potential patients who were said to have been screened for each study, and not of the (lower) number remaining following subtraction of those who were subsequently excluded because they did not meet the inclusion criteria, or for other reasons; they may thus be regarded as conservative. Since refusal to participate is a patient decision rather than one made by the study investigators, such a decision made prior to randomisation may be expected to affect both treatment groups equally, and seems unlikely to affect study validity, although it may limit generalisability if the decision to participate is influenced by symptom severity. However, in the INVEST study a comparison of data relating to eligible patients at the lead site who did and did not enrol found no significant differences in age, proportion of women, or RDQ score; data relating to pain were collected differently in the two groups and were therefore not directly comparable. The authors therefore suggested that the results of the INVEST study should be generalisable to all patients who would have been eligible for enrolment in that study.[199]

Table 7: Numbers of potential participants who refused to participate in the included studies

Study	Potential participants who refused to participate
Blasco 2012[146]	Not reported
Buchbinder 2009[101]	141/468 *(30%)*
Farrokhi [147]	2/105 (*2%*)
FREE[151]	209/1279 (*16%*)
INVEST[102]	300/1813 (*17%*)
Liu 2010[148]	Not reported
Rousing 2009[139]	Not reported
VERTOS[153]	Not clear how many potential participants were screened: the study states that, "Of approximately 1 in 4 potential study candidates, a total of 46 patients consented initially to participate in the study"
VERTOS II[17]	*277**/934 (*30%*)
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers	

* Includes 45 who requested vertebroplasty prior to randomisation

Although most of the included studies provided an explicit description of the interventions, Rousing et al only provided a relatively cursory description of the control treatment.[139] Traditionally, PVP has been performed by radiologists and BKP by surgeons.[200] However, few of the included studies provided adequate details of the clinical background and specific procedure-related training of the clinician who performed PVP or BKP, or of their relevant experience (i.e. the number of procedures they had completed before the study), although such information is important for interpreting study results. Studies involving inexperienced clinicians and centres in the early stages of introducing PVP or BKP may include "learning curve" data, whereas studies involving more experienced clinicians and centres may have more favourable results. The information provided was judged unclear if only the specialism (e.g. radiology, neurosurgery) was reported; it was only judged adequate if details were also given of the specific training in PVP or BKP which the operators had received, or the number of such procedures which they had previously performed. Thus, Buchbinder et al specified that PVP was performed by experienced interventional radiologists who had undertaken formal training in vertebroplasty, had appropriate certification, were actively performing the procedure, and all adhered strictly to a detailed, standardised protocol,[101] while in the INVEST study PVP was said to be performed by highly experienced practitioners who had performed a mean of approximately 250 procedures

(range 50-800).[102] Blasco,[146] Farrokhi,[147] and Rousing[139] provided information relating to the clinicians' specialism (respectively experienced neurointerventional radiologists, a neurosurgeon, and orthopaedic surgeons specialising in spine surgery) but did not specify their training or level of experience of PVP. The remaining studies (FREE, Liu, VERTOS, VERTOS II) provided no relevant information.

All studies used relevant outcome measures, and all but Liu specified that they used valid instruments. All assessed short-term outcomes. However, some studies did not either measure or report long-term outcomes: as noted earlier, 1-year assessments were planned in INVEST and VERTOS but, because of crossover, VERTOS was stopped at 2 weeks[153] while INVEST followed patients up for a year but only reported outcome data at 1 month.[102] Buchbinder et al followed patients up for 2 years,[101] and the 1- and 2-year data are currently being prepared for publication, as is a separate paper on radiological outcomes (Rachelle Buchbinder personal communication).

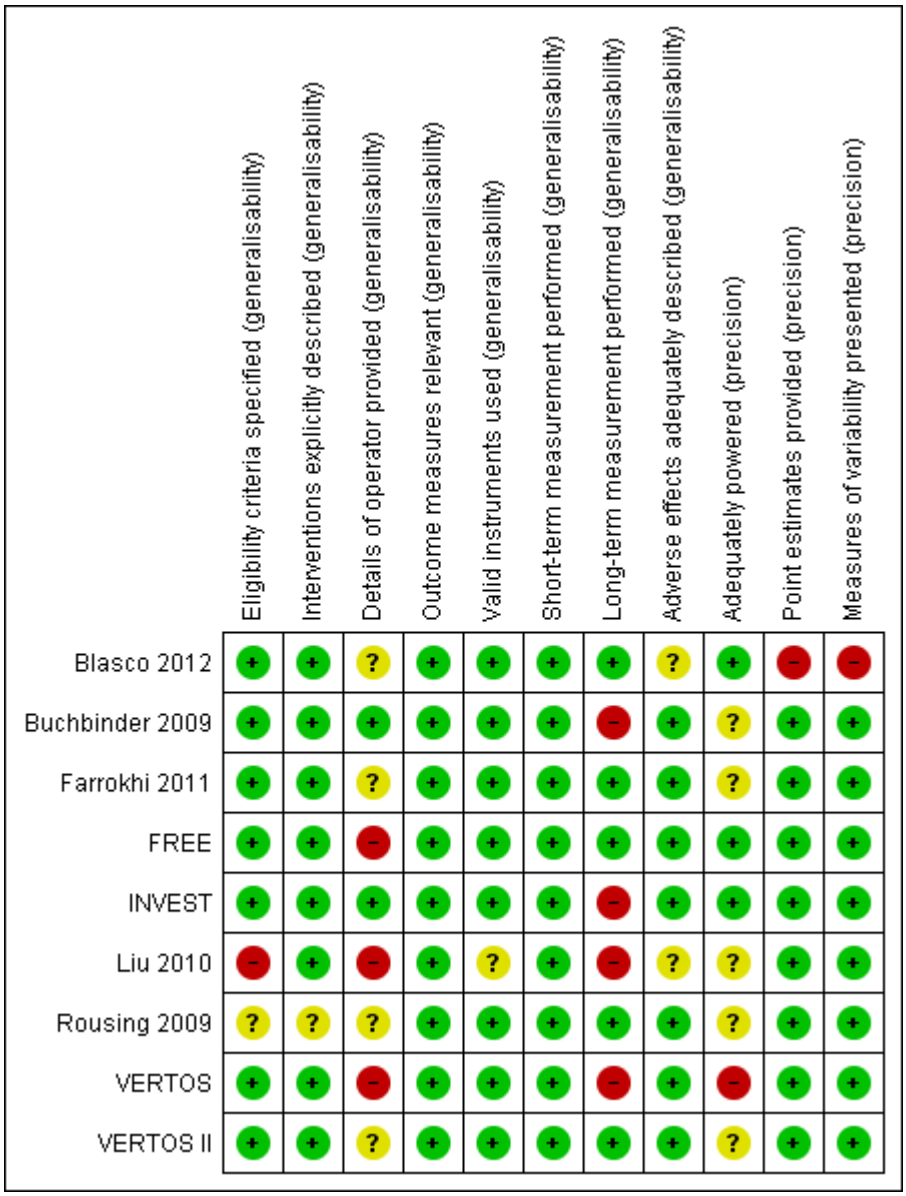
Most studies provided an adequate description of adverse effects.

Precision

Only three of the included studies (Blasco, FREE, INVEST) appeared to be adequately powered for at least their primary outcome: pain as measured on an 11-point scale (Blasco, INVEST) and change from baseline to 1 month in the SF-36 physical component summary (PCS) scale (FREE). However, because of difficulties with recruitment, the power of the INVEST study was reduced from a power of more than 80% to detect a 2.5-point difference between groups on the RDQ score and a 1.0-point difference on an 11-point pain scale to a power of more than 80% to detect a 3.0-point difference on the RDQ score and a 1.5-point difference on the pain scale.[102] Because most studies were underpowered for most outcomes, the absence of a statistically significant difference between treatment groups does not necessarily mean that such a difference would not be found in a larger study.

Almost all studies except that by Blasco et al published point estimates and measures of variability; Blasco kindly supplied additional data in that format (Andaluz Blasco personal communication).

Figure 3: External validity and precision summary: review authors' judgements about each included study (+ = good generalisability/precision; - = poor generalisability/precision; ? = unclear generalisability/precision)



Summary of internal and external validity

The quality of the included studies is generally not very high. Much of this is due to the widespread lack of blinding: the studies at least risk of bias are the double-blinded Buchbinder and INVEST studies which compare PVP with an OPLA. The studies which compare PVP with non-invasive management (Blasco, Farrokhi, Rousing, VERTOS, VERTOS II) vary in quality, that by Farrokhi et al being least at risk of bias.

The FREE study, the only study to compare BKP with non-invasive management, is at risk of bias because of the lack of blinding of patients and outcome assessors, failure to

follow up at least 80% of participants, the unexpected imbalance in drop-outs, and selective reporting of outcomes.

The only study to compare PVP with BKP, that by Liu et al, is poorly reported and potentially at risk of bias from a number of sources. It also appears to be underpowered to identify statistically significant differences in effectiveness between the two interventions.

The external validity of the included studies is limited by the fact that only two (Buchbinder, INVEST) provided adequate information about the operating clinicians' training and experience. This makes it difficult to assess to what extent study results may be replicable elsewhere. In addition, the current lack of long-term outcome data in the Buchbinder, INVEST, Liu, and VERTOS studies make it difficult to assess the value of the procedure; however, long-term data from the study by Buchbinder et al are to be published.

Only three studies (Blasco, FREE, INVEST) appeared to be adequately powered for at least their primary outcomes (pain score in Blasco and INVEST, change in SF-36 PCS score from baseline to 1 month in FREE). Because most studies were underpowered for most outcomes, the absence of a statistically significant treatment effect should not necessarily be taken as evidence that no such difference exists.

5.2.2 Assessment of effectiveness

Clinical effectiveness

Health-related quality of life

AQoL

Only one study (Buchbinder) reported AQoL scores. No difference was found between the PVP and control groups (see Appendix 9, Table 104).

Dallas Pain Questionnaire

As noted in section 4.5.1, despite its name, the DPQ was not designed to evaluate pain per se but the impact of chronic pain on various aspects of the patient's life: lower scores indicate better quality of life.[121]

Only one study that by Rousing et al, used the DPQ to evaluate PVP. Rousing et al claim that, although the other results are not statistically significant, the result for work and leisure at 3

months reaches statistical significance, favouring PVP.[139,185] This is indeed true of the unadjusted score. However, in each domain, baseline scores were noticeably lower in the PVP group than in the control group. Once this is adjusted for by comparing changes from baseline in each group rather than crude scores, it is clear that all the point estimates favour conservative management whereas previously all except that for social interest at 12 months had favoured PVP (see Table 8). Moreover, the differences are statistically significant for all outcomes except work and leisure at 3 months and anxiety and depression at both 3 and 12 months.

Table 8: Mean Dallas Pain Questionnaire scores, by domain, before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: medium- and long-term outcomes: data from Rousing et al[139,185]

Domain	Time point	PVP (95% CI)	Control (95% CI)	Mean difference between groups (95% CI) (negative values favour intervention)	P value
Daily activities	Baseline	47.8 (22.5-73.1)	68.5 (47.0-90.1)	*-20.7*	
	3 months	47.1 (32.9-61.4)	57.4 (40.7-74.1	*-10.3*	0.33
	Change from baseline at 3 months	*-0.7 (-7.75 to +6.3)*	*-11.1 (-17.24 to -4.96)*	*+10.4 (+0.83 to +19.97)*	
	12 months	53.0 (38.3-67.7)	53.6 (34.8-72.5)	*-0.6*	0.95
	Change from baseline at 12 months	*+5.2 (-1.89 to +12.92)*	*-14.9 (-21.33 to -8.47)*	*+20.1 (+10.25 to +29.95)*	
Work & leisure	Baseline	41.1 (20.7-61.5)	68.7 (47.8-89.6)	*-27.6*	
	3 months	44.5 (30.4-58.7)	65.2 (50.4-80.1)	*-20.7*	0.04
	Change from baseline at 3 months	*+3.4 (-2.28 to +9.08)*	*-3.5 (-9.29 to +2.29)*	*+6.9 (-1.31 to +15.11)*	
	12 months	46.1 (31.4-60.9)	49.2 (31.5-66.9)	*-3.1*	0.78
	Change from baseline at 12 months	*+5.0 (-0.75 to +10.75)*	*-19.5 (-25.78 to -13.22)*	*+24.5 (+15.93 to +33.07)*	
Anxiety & depression	Baseline	31.5 (12.6-50.4)	43.0 (19.9-66.1)	*-11.5*	
	3 months	28.7 (15.1-42.3)	40.0 (20.8-59.2)	*-11.3*	0.30
	Change from baseline at 3 months	*-2.8 (-9.54 to +3.94)*	*-3.0 (-11.93 to +5.92)*	*+0.2 (-11.45 to +11.85)*	
	12 months	31.3 (16.5-46.2)	35.3 (20.4-50.2)	*-4.0*	0.70
	Change from baseline at 12 months	*-0.2 (-7.30 to +6.90)*	*-4.7 (-12.81 to +3.41)*	*+4.5 (-6.72 to +15.72)*	
Social interest	Baseline	23.8 (9.9-37.7)	41.0 (23.3-58.7)	*-17.2*	
	3 months	24.1 (13.2-35.0)	30.7 (15.9-45.5)	*-6.6*	0.46
	Change from baseline at 3 months	*+0.3 (-3.77 to +4.37)*	*-10.3 (-15.17 to -5.43)*	*+10.6 (+4.07 to +17.13)*	
	12 months	32.9 (18.9-46.9)	30.7 (16.5-44.8)	*+2.2*	0.82

Domain	Time point	PVP (95% CI)	Control (95% CI)	Mean difference between groups (95% CI) (negative values favour intervention)	P value
	Change from baseline at 12 months	*+9.1 (+4.60 to +13.60)*	*-10.3 (-15.28 to -5.32)*	*+19.4 (+12.49 to +26.31)*	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

EQ-5D

Five studies (Buchbinder, FREE, INVEST, Rousing, and VERTOS II) collected EQ-5D data. However, two studies did not collect relevant data from all participants, although in both cases non-collection of EQ-5D data does not appear to be related to patient characteristics. Buchbinder et al only added this outcome to their protocol in June 2005 to allow comparison with the INVEST trial study, and therefore EQ-5D scores were only available for 30/38 participants (79%) in the intervention group and 29/40 (73%) in the control group.[101] Similarly, Rousing et al only collected EQ-5D data from November 2004, when a PhD study was affiliated to the trial: thus, scores were only available for 15/26 participants (58%) in the intervention group and 17/24 (71%) in the control group.[139] VERTOS II collected EQ-5D data throughout the study, and reported baseline values, but the investigators did not report follow-up values, although they used them to estimate quality-adjusted life years (QALYs).[17]

The two blinded RCTs (Buchbinder, INVEST) found no significant difference between PVP and conservative treatment in terms of short- or medium-term outcomes (see Appendix 9, Table 105). As Doidge et al note, the confidence intervals include effects which might favour either group, suggesting that the studies were underpowered to detect clinically important differences in this outcome.[56] However, when one-month data from the Buchbinder and INVEST studies were combined in a meta-analysis of IPD,[109] the result was not statistically significant (see Appendix 9, Table 107) and, because the MCID is 0.08,[122] the confidence interval for the pooled data only just includes the possibility of a clinically important difference favouring PVP. In the study by Rousing et al, the changes from baseline at 3 and 12 months favour conservative treatment, and suggest that the difference between groups is clinically important.

The FREE study found statistically significant differences in outcomes, favouring BKP over non-surgical management, at 1, 12, and 24 months (see Appendix 9, Tables 105 and 106). However, although at each time point the point estimate is greater than the MCID, the

confidence intervals at 3, 6, 12, and 24 months include the possibility of clinically unimportant effects.

Figure 4, 5, 6, and 7 graphically represent these change in EQ-5D observed in the Buchbinder[150], FREE[151], INVEST[152] and Rousing[185] trials respectively.

Figure 4: EQ-5D data recorded in Buchbinder et al[101]

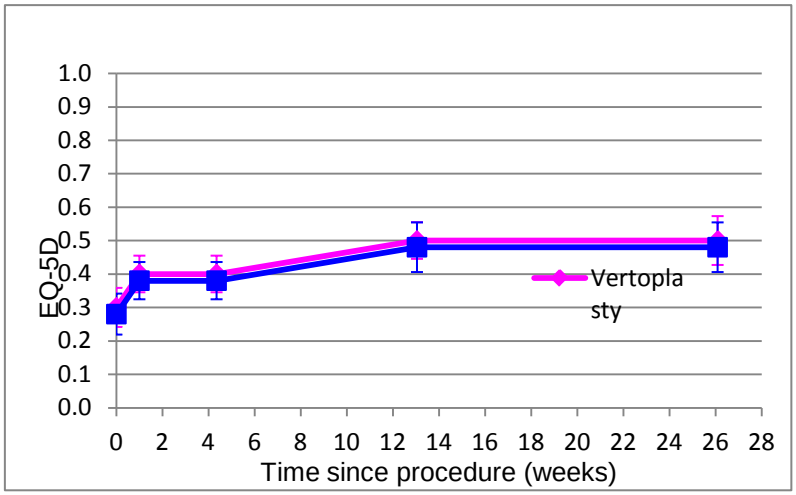


Figure 5: EQ-5D data recorded in the FREE trial[151]

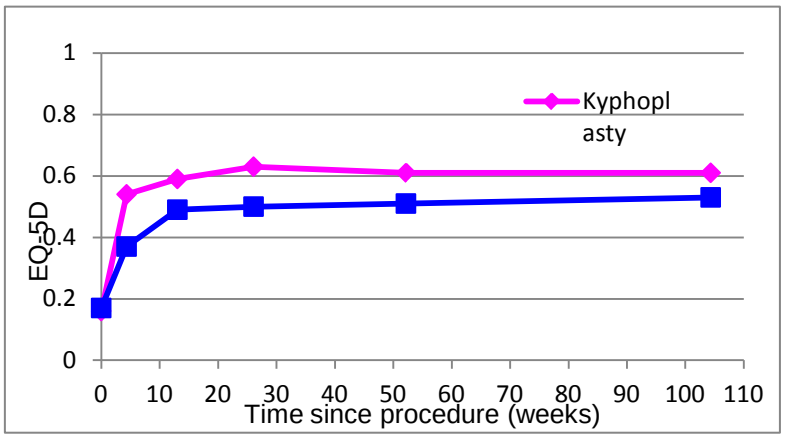


Figure 6: EQ-5D data recorded in the INVEST trial[102]

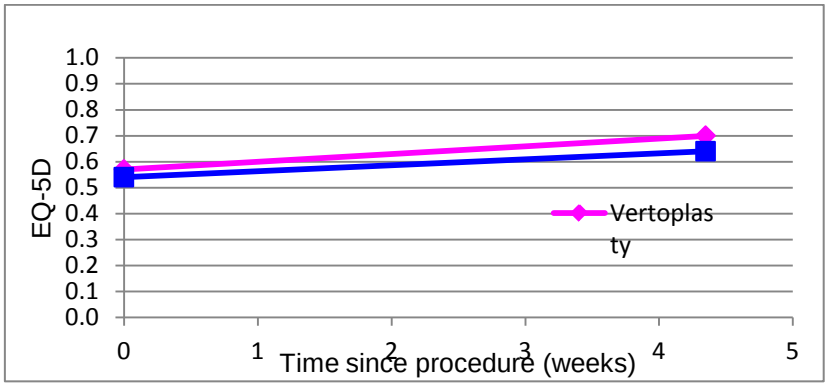
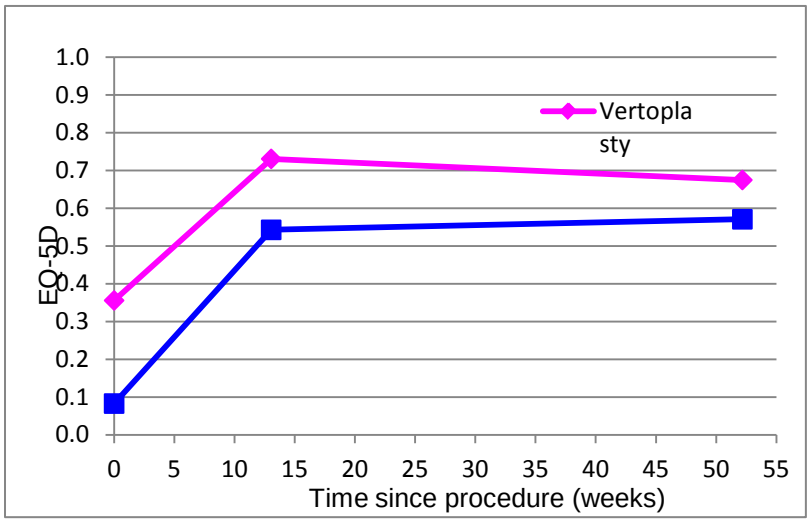


Figure 7: EQ-5D data recorded in the Rousing trial[185]



QUALEFFO

Only four studies (Blasco, Buchbinder, VERTOS, and VERTOS II) assessed health-related quality of life using the QUALEFFO (with which higher scores represent worse HRQoL). In the study by Blasco et al, although the point estimates suggest that PVP is associated with better short- and medium-term total QUALEFFO scores than conservative treatment at all time points, the confidence intervals indicate that the difference is not statistically significant (see Appendix 9, Tables 108-110). However, while Buchbinder et al stated that the only statistically significant QUALEFFO result in their study, at 1 week, favoured placebo,[101] the figures they report suggest it in fact favoured PVP (see Appendix 9, Table 108). As no MCID has been proposed for the QUALEFFO, the clinical significance of this result is not clear. The VERTOS study also found that PVP was associated with significantly better a short-term total QUALEFFO score than conservative treatment. In VERTOS II, after adjusting for baseline

differences, there was said to be a significant difference in QUALEFFO scores at 1 year which favoured PVP ($p < 0.0001$);[17] however, this result was not quantified.

SF-36

Three studies (FREE,[151] INVEST,[102] Rousing[185]) collected data relating to health-related quality of life at baseline and follow-up using the SF-36. However, only the FREE study reported mean utility scores; [redacted] (see Appendix 9, Table 111). Academic in confidence data regarding SF-36 data in the FREE study was supplied by Professor Wardlaw.

All three studies reported SF-36 PCS scores. In the FREE study, mean SF-36 PCS scores, and improvements from baseline in those scores, were reported in several publications, with some discrepancies in the results reported in the different publications: where there are discrepancies, data from the later publications have been utilised here as they are likely to be more complete. The FREE study found significant differences in medium-term outcomes; these favoured BKP. However, the between-group difference dwindled steadily from one month, when the result also suggested clinical importance: at 3 and 6 months, the confidence intervals included the possibility of failing to achieve clinical importance, while after 6 months there was no statistically significant difference between treatment groups. The INVEST and Rousing studies found no significant differences between treatment groups at any point (see Appendix 9, Tables 112 and 113).

All three studies also reported psychological wellbeing as assessed by the SF-36 MCS, and identified no statistically significant differences between treatment groups, although the confidence intervals include the possibility of potential clinically important treatment effects favouring the intervention at time points up to 6 to 12 months (see Appendix 9, Tables 114 and 115).

Back-specific functional status/mobility

All of the studies except that by Liu et al reported some measure of back-specific functional status or mobility.

RDQ

Five studies (Buchbinder, FREE, INVEST, VERTOS, VERTOS II) assessed back-specific functional status using the RDQ. Buchbinder and INVEST used the modified, 23-point, version of the RDQ; FREE used the original 24-point version as, apparently, did VERTOS

and VERTOS II. In both versions, higher scores represent worse disability; whichever version is used, the MCID appears to be at least 2 points.

Only Buchbinder, INVEST, and VERTOS reported short-term outcomes (see Appendix 9, Table 116). In terms of the between-group difference in change from baseline, all the point estimates favour PVP, but the results from the Buchbinder and INVEST studies are not statistically significant; unfortunately, the statistical significance of the result from the VERTOS study could not be calculated because of the way in which the investigators reported the results.

Buchbinder and INVEST found no significant between-group differences in medium-term outcomes. The FREE study found that BKP was associated with significantly better outcomes at 1 and 12 months, but not at 24 months; moreover, at 12 months the confidence intervals include the possibility of failing to achieve clinical importance (see Appendix 9, Table 117). VERTOS II reported a statistically significant difference in improvement over time which favoured PVP at 1 year ($p < 0.0001$); however, this was not quantified, and its clinical importance was not indicated.

Meta-analysis of IPD from the Buchbinder and INVEST studies indicated no significant difference between treatment groups at 1 month in terms of mean RDQ scores (see Appendix 9, Table 118).

In the INVEST study, a post-hoc analysis was performed to identify the proportion of patients who achieved a clinically meaningful improvement in physical disability related to back pain at 1 month: this improvement was not defined, but was presumably measured in terms of a reduction in the RDQ score. There was no significant difference between the proportion of patients in each group who achieved a clinically meaningful improvement (40% of the PVP group vs. 41% of the control group, $p=0.99$).[102] Meta-analysis of IPD from the Buchbinder and INVEST studies found no significant difference in the proportion showing an improvement of at least 3 units or of at least 30% in RDQ scores (see Appendix 9, Table 119).

Oswestry Disability Index (ODI)

None of the included studies used the original ODI. However, Farrokhi et al used a questionnaire based on the ODI which replaced the sex life dimension by a question relating to change in the degree of pain. PVP was associated with a statistically significant difference in change from baseline in the modified ODI score at all times from one week to 36 months[147]

(see Appendix 9, Table 120). Moreover, as the MCID for the ODI appears to be 4 points, these differences seem to be clinically meaningful throughout.

Barthel Index

Only one study, that by Rousing et al, reported functional outcomes using the Barthel Index, using the version scored from 0 to 20, with lower scores indicating greater disability. As data were only collected from November 2004, they are only available for a subset of the study population.[139] Rousing et al state that, at 12 months, the absolute score was significantly better in the PVP group than in the control group.[185] However, once the difference in baseline scores is taken into account, the difference between groups is no longer statistically significant (see Appendix 9, Table 121). It is difficult to know how much importance to attribute to this result as it may reflect a ceiling effect whereby, because the baseline measurement is relatively high, there is little scope for the intervention to improve the outcome beyond the extent to which it would improve under the control treatment.

SOF-ADL

The INVEST study reported mean SOF-ADL scores at baseline and 1 month. There was no statistically significant difference between treatment groups in change from baseline (see Appendix 9, Table 122).

Other indicators of disability

Three studies (Farrokhi, FREE, Rousing) provided information relating to other indicators of disability. Farrokhi et al noted that all 40 patients in the PVP group could walk one day after PVP, but only 1/42 in the control group (2%) could walk at the equivalent point in time,[147] indicating a relative risk of 28.32 (95% CI 5.88 to 136.45, $p<0.0001$).

The FREE study[151] reported the use of walking aids, back braces, miscellaneous aids, and physiotherapy: the data relating to the use of walking aids are presented in Table 9. BKP was associated with a statistically significant reduction in the risk of needing walking aids at 1 month, but not at 12 months. However, the data are not robust because, in the control group, the number of patients requiring walking aids at 12 months is smaller than the number for whom data are missing (44/107).

Table 9: Use of walking aids: data from the FREE study[151]

Time point	BKP	Control	Relative risk (95% CI)	P value
Baseline	49/148 (33%)	55/151 (36%)	*0.91 (0.67 to 1.24)*	*0.55*
1 month	33/136 (24%)	54/129 (42%)	*0.58 (0.40 to 0.83)*	*0.003*
12 months	30/121 (25%)	38/107 (36%)	*0.70 (0.47 to 1.04)*	*0.08*
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers				

The FREE study also provided data relating to the number of patients who reported one or more days of bed rest due to back pain in the previous 14 days. Again, BKP was associated with a statistically significant reduction in the risk of needing bed rest at 1 month but not at 12 months (see Table 10). However, the 12-month data in both groups are not robust because the numbers of patients for whom data are missing outnumber those who report the outcome of interest. At one month, patients in the BKP group reported on average 2.9 fewer days of restricted activity because of back pain in the previous 14 days than did controls (95% CI 1.3-4.6, p<0.001), but at 12 months the difference was no longer statistically significant (1.6 days, 95% CI -0.1 to 3.3, p=0.0678).[151] The actual numbers of days of restricted activity in each group were not reported.

Table 10: Bed rest due to back pain in the previous 14 days: data from the FREE study[151]

Time point	BKP	Control	Relative risk (95% CI)	P value
Baseline	85/146 (58%)	92/144 (64%)	*0.91 (0.76 to 1.10)*	*0.32*
1 month	30/133 (23%)	51/121 (42%)	*0.54 (0.37 to 0.78)*	*0.001*
12 months	5/120 (4%)	8/106 (8%)	*0.55 (0.19 to 1.64)*	*0.28*
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers				

Only one study, that by Rousing et al, reported three observer-assessed tests of physical function: the Tandem test, timed up & go test, and repeated chair test. Although the timed up & go test was also an outcome measure in the study by Buchbinder et al,[150] only baseline values were reported.[101] In the study by Rousing et al, data were only available for a subset of the study population. No statistically significant differences between groups were noted at 3

or 12 months[139,185] but, as baseline values were not reported, the clinical meaningfulness of this result in terms of change from baseline is not clear.

Pain/analgesic use

Pain

Only one study, the FREE study,[151] reported pain using the recommended measure of global pain severity, the bodily pain subscale of the SF-36,[141] in which higher scores represent better health. The only result which has been published from this study using this measure is the difference between treatment groups in average improvement over a period of 12 months: this was 9.2 points greater in the BKP group than in the control group (95% CI 3.9-14.6, p=0.0008).[151] Fuller confidential data presented in supplementary document 8 of the manufacturer's submission[34] indicate that, [redacted] (see Appendix 9, Table 112). [redacted]

All nine studies reported pain measured on either a numeric rating scale or visual analogue scale, with higher scores indicating more severe pain. Farrokhi and INVEST asked patients to report their average pain over the previous 24 hours, while Buchbinder and FREE asked them to do so over the previous week , and the remaining studies did not specify the time period. However, empirical data indicate that broadly comparable results are obtained regardless of whether patients are asked to report average pain over the previous 24 hours or the previous week.[201] Academic in confidence data were provided for the Buchbinder RCT (Margaret Staples personal communication) for the VAS scores at 12 and 24 months.

As noted in section 4.1, the VAS is less responsive than the numeric rating scale; this is presumably the reason why Doidge et al have suggested that data collected by the two methods should not be combined in a meta-analysis.[56] The majority of the included studies (Buchbinder, Farrokhi, FREE, INVEST, Liu, VERTOS, and VERTOS II) clearly used a numeric rating scale: although some termed it a VAS, they also referred to it as a 10-point scale. It is not wholly clear whether Blasco et al actually used a VAS, although they claimed to do so. Rousing et al specified that they used a 10cm VAS,[139] and presumably did so at most time points, but they clearly used a numeric scale in a supplementary telephone interview in

which, after all but three had completed 12 months' follow-up, patients were asked to rate their back pain one month after discharge from hospital on a scale of 0-10.[185] As Doidge et al have pointed out,[56] because these data were collected almost a year after the event, they are at high risk of recall bias.

Farrokhi, FREE, Rousing, VERTOS II found statistically significant differences between groups in short- and medium-term changes from baseline in pain following PVP or BKP; FREE and VERTOS II also found statistically significant long-term differences between groups (see Appendix 9, Tables 124-126). However, the double-blinded studies (Buchbinder, INVEST), and the small VERTOS study, found no significant differences between treatment groups, while in the study by Blasco et al statistical significance in change from baseline was only reported at 2 months, when the result favoured PVP. There appears to have been no significant difference between treatment groups in terms of change from baseline at 12 months, and Blasco et al attribute the similar prevalence of moderate and/or severe residual pain to the more frequent use of rescue therapy in the control group and the higher number of new clinical fractures associated with PVP in the intervention group.[146] Moreover, the favourable result reported by Rousing et al at one month (see Appendix 9, Table 125) is unreliable because of the high risk of recall bias discussed above. Liu et al found no significant differences between PVP and BKP, but the study does not appear to have been powered to do so.

Meta-analysis of IPD from the Buchbinder and INVEST studies again found no significant difference at 1 month in terms of mean pain scores (see Appendix 9, Table 127).

A comparison of longitudinal trends in pain reduction between the differently treated groups proves instructive. Figures 8, to 13 graphically represent these trends for PVP, BKP, OPLA, and conservative treatment respectively. Graphs for the longitudinal pain changes in individual trials are also included in Appendix 11. Among the cohorts treated with PVP and BKP, there is a rapid post procedural reduction in pain which appears to stabilise at approximately one month. The OPLA treated cohorts reveal a somewhat similar pattern: there is a rapid reduction in pain, which seems to stabilise at one month. In contrast to PVP and BKP, however, there appears to be a small, temporary worsening of pain between one day and one month, at which point pain once again reduces and stabilises. A rather different pattern emerges with respect to those treated with OPM. There is no dramatic initial drop in pain; rather, there is a more gradual reduction until approximately three months, at which point the pain level stabilises and becomes comparable to those treated with PVP.

The gradual reduction in pain seen in conservatively treated patients coheres with a regression to the mean as would be expected from the natural history of healing in osteoporotic VCFs. The patterns seen in the PVP and OPLA groups, on the other hand, pose some more interesting interpretive questions. Whitehouse has suggested that the initial 'dip' seen in the OPLA cohorts represents a strong initial placebo effect before regression to the mean, while the early and sustained reductions in the PVP cohorts is suggestive of specific mechanisms of efficacy.[94] However, due to the truncated line from INVEST, interpretations should be made with caution.

Figure 8: Longitudinal pain reduction trends in vertebroplasty without AIC data.

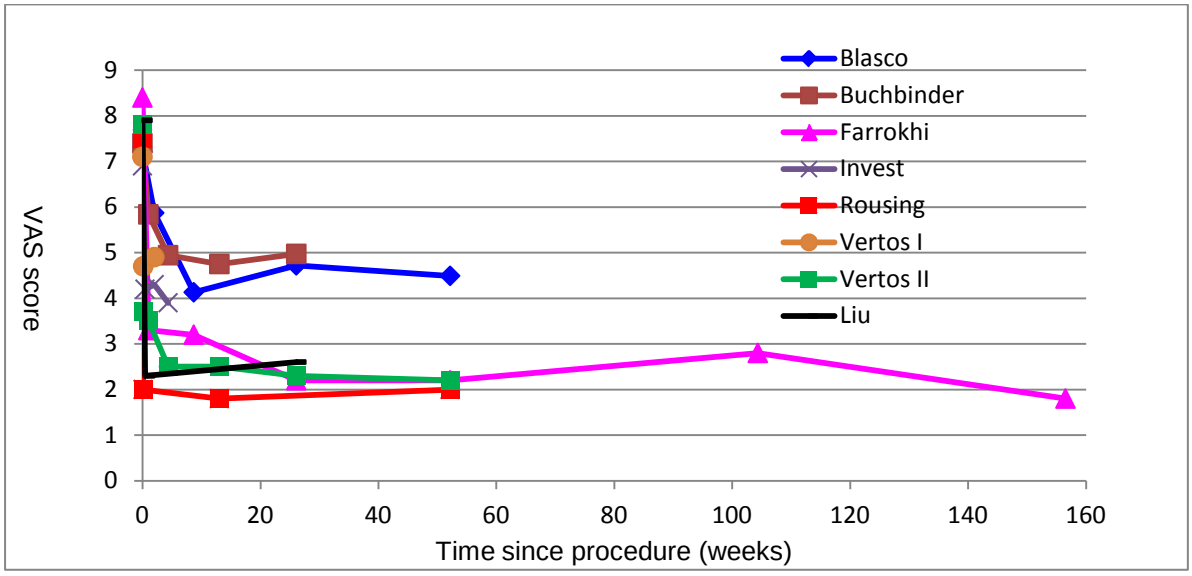


Figure 9: Longitudinal pain reduction trends in vertebroplasty with AIC data.

[redacted]

Figure 10: Longitudinal pain reduction trends in balloon kyphoplasty

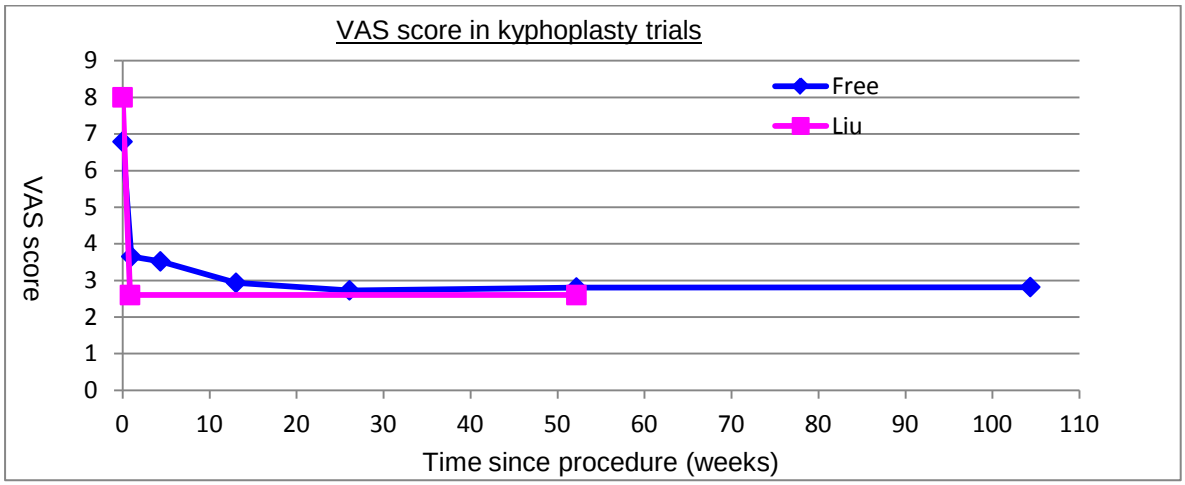


Figure 11: Longitudinal pain reduction trends in OPLA excluding AIC data.

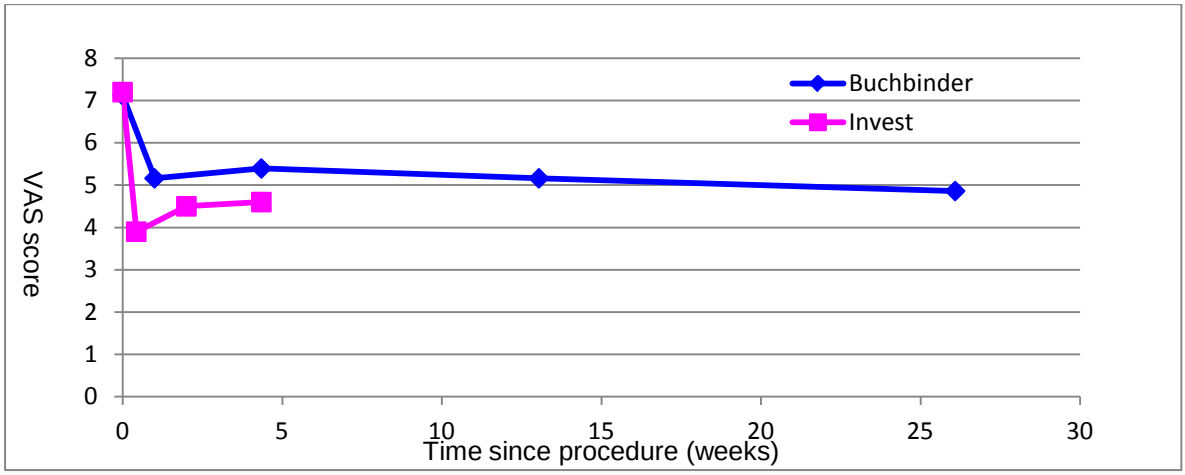
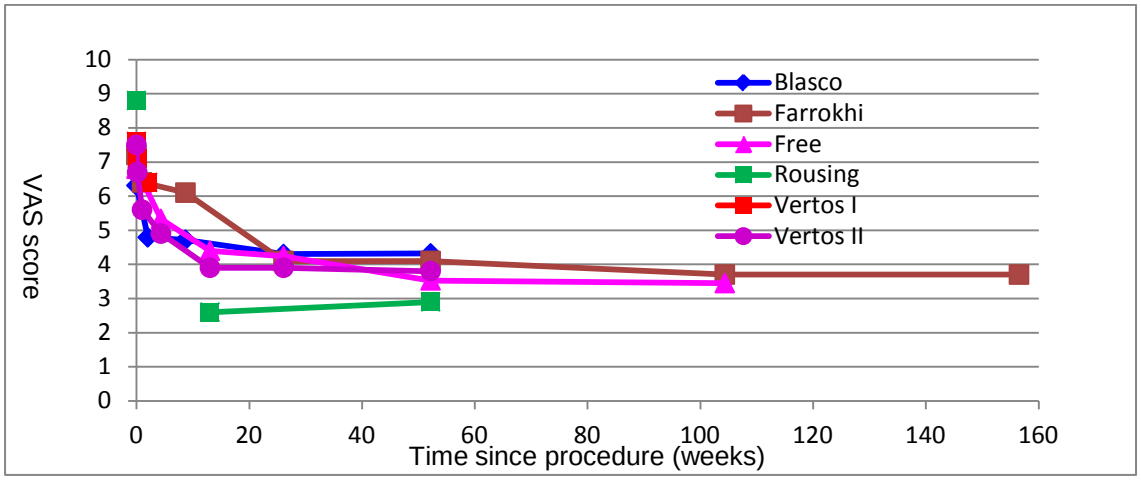


Figure 12: Longitudinal pain reduction trends in OPLA including AIC data.

[redacted]

Figure 13: Longitudinal pain reduction trends in OPM.



If, as indicated in section 4.1, a difference between groups of 2 or more points indicates a clinically meaningful difference, most of the short- to medium-term results which are statistically significant also appear to be clinically meaningful. However, although FREE and VERTOS II both reported statistically significant longer-term results, in the FREE study these results did not appear to be, and in VERTOS II the 95% CI included the possibility that they were not, clinically meaningful (see Appendix 9, Table 126).

The INVEST study also stated that 64% of patients randomised to PVP and 48% of those randomised to OPLA reported a clinically meaningful improvement in pain (i.e. a decrease of 30% or more) at 1 month (p=0.06).[102] It has been suggested that this trend towards favouring PVP might have achieved statistical significance if the trial had recruited more participants, as was originally planned.[202] However, when these data were combined with those from the Buchbinder study in a meta-analysis of IPD, no significant difference at the 5% level was found in the proportion showing a clinically meaningful improvement in pain at 1 month, whether this was defined as a decrease in pain of at least 3 units or of at least 30% (see Table 11).

Table 11: Number of patients in the Buchbinder and INVEST studies showing improvement in pain scores at one month: data from Staples et al 2011[109]

Outcome	PVP	Control	Relative risk (95% CI)	P value
Improvement in pain of ≥3 units	55/102 (53.9%)	43/99 (43.4%)	1.3 (0.8 to 1.9)	NS
Improvement in pain of ≥30%	61/102 (59.8%)	45/99 (45.5%)	1.3 (1.0 to 1.8)	NS

In VERTOS II, survival analysis indicated that significant pain relief (apparently defined as a decrease from baseline in pain score of 3 points or more) was achieved earlier, and in more patients, after PVP than with conservative treatment (29.7 days (11.45-47.97) vs. 115.6 days (85.87-145.40) (χ^2=55.6, p<0.0001)).[17]

Blasco, Buchbinder, and VERTOS also reported pain outcomes in terms of QUALEFFO pain scores. The reported figures are not directly comparable as they appear to use different scales, although this is poorly reported: Blasco appears to report the domain score (scored from 0 to 5) whereas Buchbinder reports pain scores on a scale of 0 to 100, and it is not clear what potential range of scores is represented by the VERTOS data. Blasco and Buchbinder found no statistically significant difference between the groups; the significance of the VERTOS results unfortunately could not be calculated (see Appendix 9, Table 128).

The INVEST study also reported on the frequency with which participants experienced pain, and the impact of pain on their daily lives, both measured on a scale of 0-4. In both groups, pain frequency and pain bothersomeness decreased between baseline and one month; however, although the point estimates favoured the intervention, the results were was not statistically significant[102] (see Appendix 9, Table 129). Moreover, as Doidge et al note, the confidence intervals did not include a 1-unit effect (the smallest possible threshold of clinical importance) in either direction.[56]

Buchbinder et al collected data on perceived pain: this was classified as "better" if the patient indicated that it was moderately or a great deal better than before the intervention, and "worse" if they reported that it was moderately or a great deal worse. They found no statistically significant between-group differences in the proportion of patients in these categories at any time-point[101] (see Appendix 9, Table 130).

Analgesic use

Six studies (Blasco, Buchbinder, FREE, INVEST, VERTOS, VERTOS II) reported analgesic use. Blasco et al divided analgesic use into four categories: no treatment; minor analgesics (paracetamol and/or NSAIDs); minor opiate derivatives; and major opiate derivatives. They found no significant changes between groups in the analgesia used throughout the study (χ-square test, adjusted p-values >0.05)[146] (see Appendix 9, Table 131). However, rescue therapy by intrathecal infusion of 25 μg fentanyl and 1.5 mg bupivacaine was offered to patients in either group with a pain score of 7 or over; over the 12-month study period, and was required by substantially more patients in the control group than in the PVP group (15/61 (25%) compared with 3/64 (5%), p=0.0015),[146] suggesting greater pain in the control group.

Three studies (Buchbinder, FREE, and INVEST) reported the number of patients in each group who took opioids for pain at baseline and at follow-up. For comparability with these studies, Blasco et al's data on minor and major opiate derivatives were pooled to produce a total number of patients taking opioids. Review Manager was then used to calculate the relative risks of taking opioids for each of the four studies; for the INVEST study, the numerator in each group was inferred at one month from the denominator and proportion. In the Buchbinder and INVEST studies, the number of patients taking opioids for pain decreased over time in both the PVP and control group; in the Blasco study, no significant between-group differences were observed other than at baseline. However, in the FREE Study, BKP was associated with a significantly reduced risk of requiring opioid medication at 1 month and 6 months, but not at 12 or 24 months. The results from the study by Blasco et al are difficult to interpret. This is partly because a statistically significantly higher proportion of participants in the PVP group required opioid analgesia at baseline, but the picture thereafter is puzzling: in the PVP group, the proportion of participants requiring opioid medication falls noticeably from baseline to 2 weeks, and then gradually thereafter, as might be expected, whereas in the control group it rises steeply at 2 weeks and remains elevated for 6 months, then falling substantially at 12 months (see Table 12). However, in this study, in both treatment groups the number of patients requiring opioid analgesia at 12 months is smaller than the number for whom data are missing (23/64 randomised to PVP and 19/61 randomised to control), and therefore the data are not robust.

Table 12: Number of patients using opioids before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	BKP	Control	RR (95% CI)	P value
Blasco et al[146]	Baseline	*47/64 (73%)*	N/A	*31/60 (52%)*	*1.45 (1.08 to 1.93)*	*0.01*
	2 weeks	*33/56 (59%)*	N/A	*36/58 (62%)*	*0.95 (0.71 to 1.28)*	*0.73*
	2 months	*30/52 (58%)*	N/A	*33/56 (60%)*	*0.98 (0.71 to 1.35)*	*0.90*
	6 months	*26/49 (53%)*	N/A	*31/52 (60%)*	*0.89 (0.63 to 1.26)*	*0.51*
	12 months	*22/41 (54%)*	N/A	*17/42 (40%)*	*1.33 (0.83 to 2.11)*	*0.23*
Buchbinder 2009[101]	Baseline	30/38 (79%)	N/A	34/40 (85%)	*0.93 (0.75 to 1.15)*	*0.49*
	1 week	*27/38 (71%)*	N/A	*27/40 (68%)*	*1.05 (0.78 to 1.41)*	*0.73*
	1 month	*26/38 (68%)*	N/A	*25/40 (63%)*	*1.30 (0.5 to 3.32)*	*0.58*
	3 months	*19/38 (50%)*	N/A	*23/40 (58%)*	*0.74 (0.30 to 1.81)*	*0.51*
	6 months	*13/38 (34%)*	N/A	*16/40 (40%)*	*0.78 (0.31 to 1.96)*	*0.60*
FREE[151,196]	Baseline	N/A	103/140 (73.6%)	99/146 (67.8%)	*1.08 (0.93 to 1.26)*	*0.28*
	1 month	N/A	53/114 (46%)	74/115 (64%)	*0.48 (0.28 to 0.82)*	*0.007*
	6 months	N/A	37/124 (29.8%)	48/112 (42.9%)	*0.70 (0.49 to 0.98)*	*0.04*
	12 months	N/A	33/118 (28.0%)	34/101 (33.7%)	*0.83 (0.56 to 1.24)*	*0.36*
	24 months	N/A	10/114 (8.8%)	10/105 (9.5%)	*0.92 (0.40 to 2.12)*	*0.85*
INVEST[102]	Baseline	38/68 (56%)	N/A	40/63 (63%)	*0.88 (0.66 to 1.17)*	*0.38*
	1 month	*36*/67 (54%)	N/A	*26*/61 (43%)	*1.26 (0.78 to 1.82)*	*0.22*

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

In their meta-analysis of IPD from the Buchbinder and INVEST studies, Staples et al[109] found that, after adjusting for baseline opioid use, patients randomised to PVP were more likely to be taking opioids at one month than patients randomised to placebo (RR 1.25, 95% CI 1.14-1.36, p<0.001).[109] They consequently suggest that the trend observed in their meta-analysis, towards a higher proportion of patients in the PVP group achieving an improvement of 30% or more in pain scores at 1 month, may have been influenced by the fact that the PVP group was more likely than the placebo group to be using opioids at that point.[109] This contrasts with the FREE study in which opioid use was similar in both groups at baseline, but the BKP group was significantly less likely than the control group to be using opioid analgesia at 1 and 6 months (see Table 12), while the reduction in pain at 1 month was significantly greater in the BKP group than in the control group; the significance of the result at 6 months is not clear (see Appendix 9, Table 125).

Data from VERTOS and VERTOS II are not comparable with data from the four studies reported above. VERTOS recorded opioid use at baseline, but data at 1 day and 2 weeks were only reported in terms of a mean analgesic use score derived by classifying no medication as 0, paracetamol as 1, NSAIDs as 2, and opiate derivatives as 3. There was said to be no significant between-group difference in the use of pain medications at baseline (p=0.5). However, at both 1 day and 2 weeks, the mean analgesic use score had reduced in the PVP group and increased in the control group,[153] resulting in statistically significant differences which favoured PVP (see Appendix 9, Table 132). At the same points in time, there were significantly greater reductions in pain in the PVP group than in the control group (see Appendix 9, Table 124), and thus pain and analgesic use had reduced in parallel.

In VERTOS II, the class of drugs used for pain relief was said to be similar in both groups at baseline (see Appendix 9, Table 106): although 53% of patients in the PVP group used either weak or strong opioid derivatives, compared with 46% in the control group, this difference was not statistically significant (*p=0.34*). Analgesic use was said to be significantly reduced in the PVP group compared with the control group at 1 day, 1 week, and 1 month (p<0.0001, <0.001, and 0.033 respectively), but not at later stages of follow-up;[17] however, the actual figures were not presented. Pain scores were also lower in the PVP group than in the control group at 1 day, 1 week, and 1,3, 6, and 12 months, though the between-group difference in change from baseline was only said to be statistically significant at 1 and 12 months (see Appendix 9, Tables 125 and 126).

Vertebral body height and angular deformity

Four studies (Blasco, Farrokhi, FREE, Liu) reported changes in VBH and/or angular deformity. However, their results are not necessarily comparable since it is not clear that they used the same methods of measuring vertebral height. Farrokhi et al[147] calculated the mean VBH by taking the mean of the height of the anterior wall plus the height of the posterior wall, while Blasco et al[146] and Liu et al[148] referred to the mean height without specifying how it was measured. Farrokhi et al specified that they used the sagittal index (SI) to measure angular deformity[147] whereas the FREE study[203] and Liu et al[148] used the kyphotic angle (see section 4.1). It is not clear whether Liu et al[148] measured post-operative VBH and angular deformity at 3 days or at 6 months. Because of these potential sources of heterogeneity, it did not seem appropriate to pool data relating to VBH or angular deformity.

Surprisingly, the FREE study did not report changes in VBH even though maintenance of VBH was one of its secondary outcome measures,[204] and is one of the respects in which BKP might be expected to provide additional benefit compared with PVP. However, although kyphotic angle was measured in both groups, the study protocol stated that VBH was only to be measured in patients undergoing BKP,[204] thus making comparison with controls impossible.

Blasco et al found no significant difference between treatment groups in change in VBH from baseline at 12 months. By contrast, Farrokhi et al found that PVP was associated with significant improvements in mean VBH which were sustained throughout the first year but not thereafter, and with significant improvements in angular deformity which were sustained throughout the 36-month follow-up period (see Appendix 9, Tables 134 and 135). They suggest that the significant differences from pre-treatment values in mean VBH and SI seen at one week in the PVP group ($p<0.002$ and <0.011 respectively) but not in the control group ($p=0.22$ and <0.80 respectively) may be related either to the prone position used during PVP or to the high pressure produced within the vertebra by the injected cement, both of which can expand the vertebra and correct kyphotic deformity to some extent.[147]

The FREE study only reported improvement from baseline in the kyphotic angle of the index fracture at 24 months, without reporting the absolute figures at either point in time. They reported a statistically significant result in favour of BKP;[203] however the clinical significance of this result is not clear. In the study by Liu et al, BKP was associated with significantly greater improvements in both postoperative VBH and angular deformity than was PVP (see Appendix 9, Tables 134 and 135).

Progression of treated fractures

Only one study, VERTOS II, reported data relating to the progression of treated fractures during follow-up; in the control group, all vertebrae which showed bone oedema on baseline MRI were considered to be treated vertebrae. At last follow-up (mean 11.4 months, median 12.0 months, range 1-24 months), moderate or severe height loss was seen in 11 vertebrae in 11/91 patients (12%) in the PVP group, compared with 39 vertebrae in 35/85 patients (41%) in the control group ($p<0.001$)[165] (see Appendix 9, Table 136).

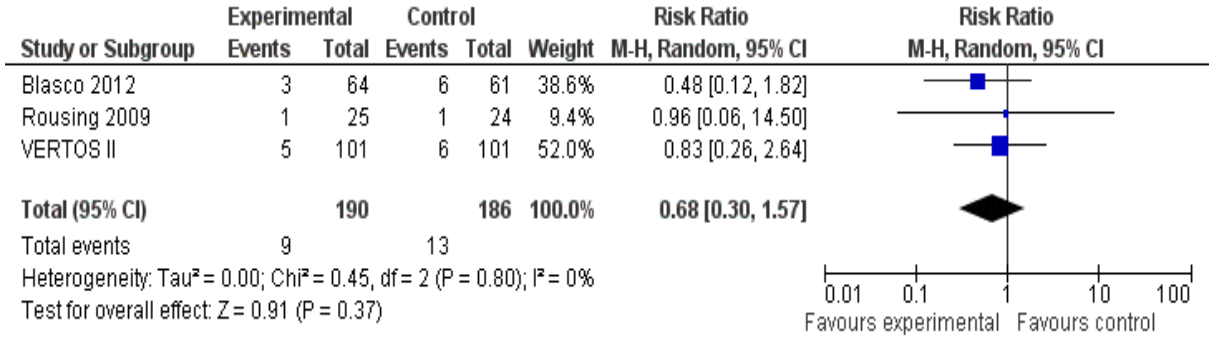
Adverse effects

All-cause mortality

Seven of the included studies reported all-cause mortality. Liu et al[148] made no reference to any deaths, thus implying that none occurred; however, this was not explicitly stated. None of the individual studies found any statistically significant differences in overall mortality between treatment groups (see Appendix 9, Table 137). However, this is unsurprising as they were not powered for this outcome. None of the reported deaths appear to be related to treatment: the patient in VERTOS II who died as a result of gastric bleeding had used morphine as their only analgesic.[17]

Three studies (Blasco, Rousing, VERTOS II) reported overall mortality at the same time point (12 months). Data from these studies were combined by meta-analysis; inclusion of data from other studies which reported mortality at different time points was not considered appropriate. Statistical significance was still not achieved when the data from these studies were pooled, although the point estimate favours PVP (see Figure 14).

Figure 14: Overall mortality at 12 months



Symptomatic and asymptomatic cement leakage

Seven studies (Blasco, Buchbinder, Farrokhi, FREE, Rousing, VERTOS, VERTOS II) reported cement leakages identified using imaging equipment. Four studies (Blasco, Buchbinder, Farrokhi, Rousing) appear to have reported only leakages identified by fluoroscopy during the procedure, whereas two (VERTOS, VERTOS II) performed a CT scan immediately after PV to identify possible cement leakage or other local complications; this technique is likely to identify more leaks than fluoroscopy. The FREE study assessed cement extravasation using both intraoperative fluoroscopy and postoperative radiographs,[151] thus increasing their likelihood of identifying leaks compared with the use of fluoroscopy alone (see section 4.1). All seven studies stated that they used PMMA cement; none referred specifically to high-viscosity cement, and it is therefore assumed that low-viscosity cement was used in all studies.

For PVP, the number of treated vertebrae in which cement leakages were reported ranged from none in the small VERTOS study to 72% in VERTOS II; the pooled data suggest an incidence of 44% for PVP compared with 27% for BKP. However, this approach may conceal a relationship between the volume of cement injected and the likelihood of leakage, or between the sensitivity of the method of detection used and the detection rate: thus, the highest incidence is seen in VERTOS II, which also reports the highest mean volume of cement injected per vertebra, and which specifically scanned patients post-operatively using CT scanning, the most sensitive method of detection, to identify possible leakages (see Table 13).

Table 13: Number of treated vertebrae with imaging-identified cement leakage

Study	**Mean (SD) volume of cement injected (per vertebra)** (ml)	**PVP**	**BKP**
Blasco 2012[146]	NR	*67*/140 (49%, *95% CI 41 to 57%*)	
Buchbinder 2009[101]	2.8 (1.2)	NR†	
Farrokhi 2011[147]	3.5 (range 1-5.5)	14/100 (14%, *95% CI 7% to 21%*)	
FREE[151]	NR		51/188†† (27%, *95% CI 21% to 33%*)
Rousing 2009[139]	NR	NR*	
VERTOS[153]	3.2 (range 1-5)	0/*29*	
VERTOS II[17]	4.1 (1.5, range 1-9)	97/134 (72%, *95% CI 64% to 80%*)	
Total		*178/403* *(44%, 95% CI 39% to 49%)*	51/188 (27%, *95% CI 21% to 33%*)
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers			

† Minimal leakage was recorded in 14/38 patients (37%);[101] the number of affected vertebrae was not reported; †† 32% of patients (48/149) were affected[151]; * Extravertebral leaks were said to have occurred, but the number was not reported.[139]

The importance of cement leaks relates to their potential clinical sequelae. These may be immediate or delayed. Blasco found that, although the cement leaks which they reported were not associated with immediate clinical complications, cement leakage into the inferior disk was associated with an increased risk of incident vertebral fracture (OR 7.17 (1.69-69.30), p=0.0008).[146] Farrokhi reported 13 asymptomatic leaks (5 into the discal space and 8 into the paravertebral space), and one symptomatic leakage into the epidural space. The symptomatic leakage caused severe right lower-extremity pain and weakness but, following immediate decompression through a bilateral laminectomy and evacuation of bone cement, the patient could walk unassisted with no radicular pain after 2 months.[147] Rousing stated that none of the cement leaks caused neurological symptoms.[139] In VERTOS II, most leakages were discal or into segmental veins; none were into the spinal canal. All patients remained asymptomatic, even though fluoroscopy showed cement migration into the venous system towards the lungs in one patient; a follow-up chest CT after 1 year showed no perifocal inflammatory pulmonary changes in this patient. In this study, an asymptomatic cement deposition in a

segmental pulmonary artery was also reported, presumably in another patient.[17] 54 PVP patients subsequently underwent CT after a mean follow-up of 22 months (median 21 months, range 6-42 months). Although, during the procedure, the operators had not reported fluoroscopically-visible cement migration towards the lungs in any of these patients, at follow-up 14/54 (26%, 95% CI 16-39%) had pulmonary cement embolism (PCE) visible on CT. The emboli varied in size between 1 and 12 mm and were randomly distributed in the periphery of the lungs; 6 patients had a single cement embolus, while the remaining 8 had between 2 and 35 cement depositions randomly scattered in the peripheral portions of both lungs. All the affected patients were asymptomatic.[172] In the FREE study, most leaks were endplate or discal leakages, with one foraminal leakage, no leakages to the spinal canal, and no cement embolisms.[151]

Periprocedural balloon rupture

Neither of the studies of BKP reported periprocedural balloon rupture.

Peri- and post-operative complications (including infection)

Seven studies (Buchbinder, Farrokhi, FREE, INVEST, Rousing, VERTOS, VERTOS II) provided some information relating to peri- or post-operative complications.

Intra-operative complications

In the INVEST study, one patient had an injury to the thecal sac during PVP which resulted in hospitalisation. In addition, one patient who had received OPLA was hospitalised overnight after the procedure with tachycardia and rigors of unknown cause.[102]

In the VERTOS study, in a patient originally randomised to optimum pain medication who requested PVP after two weeks, an intrapedicular cement spur broke on manipulation by the bone biopsy needle and caused a small cortical chip fracture at the medial border of the pedicle. The patient recorded an increase in pain score at 1 day but the pain was relieved using analgesics and local anaesthetic infiltration of the involved pedicle; there were no neurological sequelae.[153]

In VERTOS II, the patients required additional intravenous analgesia in 30% of procedures (31/98); two patients needed atropine because of pain-induced vasovagal reactions. In one case, the procedure had to be stopped because the patient developed an acute asthma exacerbation during vertebroplasty; the procedure was performed successfully a week later.[17]

Rousing et al stated that no conversions to open surgery were necessary in their study.[139] While this is not specified in relation to any of the other included studies, it seems likely that, had such conversions been required, they would have been reported.

Post-operative complications

Three studies (Buchbinder, FREE, VERTOS II) reported post-operative infections which were potentially related to treatment. Farrokhi et al specified that no infections occurred,[147] while Rousing indicates this by stating that there were no adverse events other than cement leaks.[139] In the remaining four studies (Blasco, INVEST, Liu, VERTOS) no post-operative infections were mentioned, again suggesting that none may have occurred. In the study by Buchbinder et al, prophylactic cephalothin was usually administered intravenously immediately after cement injection.[150] Osteomyelitis developed in a patient who did not receive such prophylaxis because of multiple drug allergies; surgical drainage and antibiotic treatment were required approximately two weeks after randomisation, and the patient then recovered fully.[101] In the FREE study, a recurrent urinary tract infection (UTI) was exacerbated by catheterisation; this patient also developed spondylitis near the cement in the vertebral body 376 days after surgery and the inflammation had not resolved by 24 months despite antibiotic therapy.[151] Sepsis/septic shock was reported in one patient in the BKP group, but also in three patients in the conservative treatment group.[196] In VERTOS II, one patient developed a UTI after vertebroplasty.[17]

Wardlaw et al noted that, in the FREE study, three patients who underwent BKP subsequently had pulmonary embolisms; the earliest of these was 46 days postoperatively.[151] The significance of these embolisms is not discussed.

Incidence of new vertebral fractures

Radiographic fractures

Only three studies (Blasco, FREE, VERTOS II) reported the number of patients who suffered new radiographic vertebral fractures during the study period. None of these studies found a statistically significant difference between treatment groups (see Table 14). However, in the FREE study, loss to follow-up is higher in the control group than in the BKP group (34/149 (23%) vs. 56/151 (37%)); as the drop-out rate outnumbers the event rate in the control group, the fracture incidence data may be biased.

Rousing et al only reported the number of new radiographic fractures, rather than the number of patients who suffered such fractures, and therefore the relative risk of fracture cannot be

calculated. They found that, over 12 months, there were more radiographic fractures in the PVP group than in the control group (7 vs. 4, statistical significance not reported).[185]

Although the study protocols for the Buchbinder and INVEST studies specified the incidence of new vertebral fractures as an outcome[150,152] the relevant results have not yet been published.

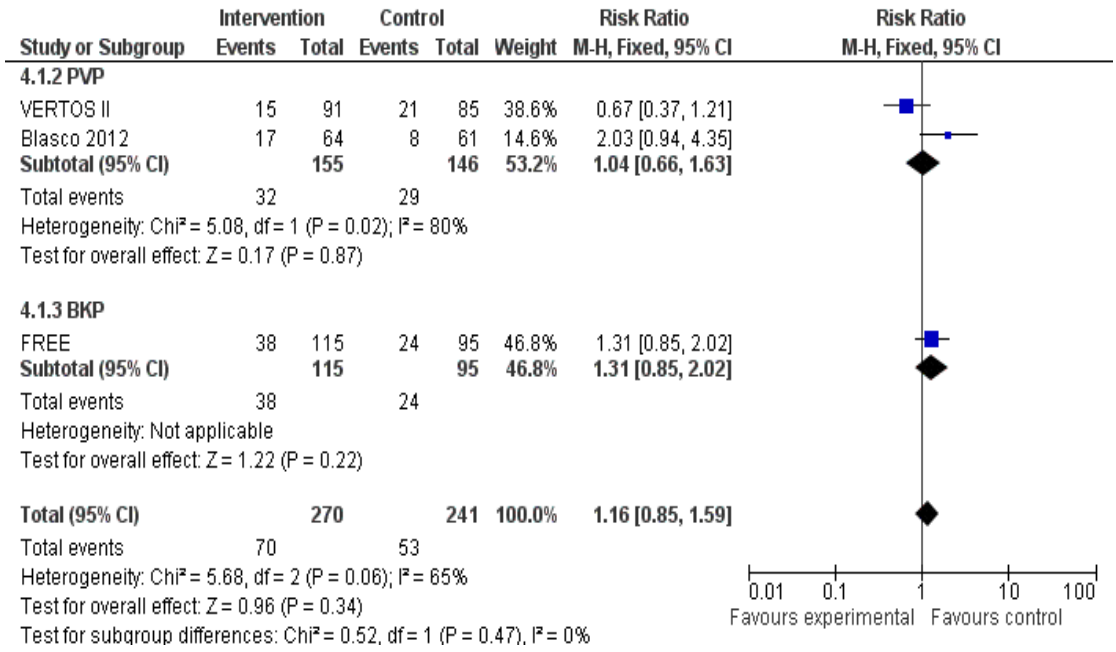
Table 14: Number of patients suffering new incident radiographic vertebral fractures

Study	**Length of follow-up**	**Mean time from estimated fracture onset to intervention (weeks)**	**No of patients with incident VCF**			**RR (95% CI)**	**P value**
			PVP	**BKP**	**Control**		
Blasco 2012[146]	12 months	*20.0 (13.7)/) 20.4 (18.6)*	17/64 *(26.6%)*	N/A	8/61 *(13.1%)*	*2.03 (0.94 to 4.35)*	*0.07*
FREE[151,196]	12 months*	5.6 (4.4)/ 6.4 (5.2)	N/A	38/115 (33%)	24/95 (25%)	*1.31 (0.85 to 2.02)*	*0.22*
	24 months		N/A	56/118 (47.5%)	45/102 (44.1%)	*1.08 (0.81 to 1.44)*	*0.62*
VERTOS II)[17]	12 months†	*4.2 (2.4)/ 3.8 (2.3)*	15/91 (16.5%)	N/A	21/85 (24.7%)	*0.67 (0.37 to 1.21)*	*0.18*
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers							

* Includes new and worsening fractures; † Mean follow-up 11.4 months (median 12, range 1-24)

As noted in section 5, data from populations with differences in length of follow-up are not directly comparable. However, as all three studies reported results at 12 months, we have performed an exploratory meta-analysis combining data from the three studies which reported the number of patients who had suffered new radiographic vertebral fractures by that time. Although the point estimate favours control, statistical significance was not achieved (see Figure 15).

Figure 15: Patients with new incident radiographic vertebral fractures at 12 months



It has been observed that vertebrae adjacent to those treated with PVP or BKP may be particularly susceptible to subsequent fractures (detailed later). Thus, fractures in adjacent vertebrae are more likely to be associated with therapy than fractures in more distant vertebrae. Blasco et al found that 82% of new fractures in the PVP group were adjacent to the index vertebra, compared with 27% in the control group (OR 16.00, 95% CI 1.03 to 835.12, p=0.0101).[146] The FREE study reported that 28/118 patients in the BKP group (23.7%) and 17/102 in the control group (16.7%) suffered a radiographic fracture adjacent to the index fracture;[196] however, the difference was not statistically significant (RR *1.42, 95% CI 0.83 to 2.45, p=0.20*). Similarly, Klazen et al reported that, in VERTOS II, the risk of adjacent rather than distant fracture was not significantly different in the intervention and control groups (p=0.23), nor did such fractures occur significantly sooner in the PVP group than in the conservative therapy group (4.6±5.4 vs. 6.1±5.9 months, p=0.48). The only risk factor for either the occurrence or the number of new fractures was the number of vertebral fractures at study entry, which is itself an indicator of the severity of osteoporosis.[165]

Clinical fractures

As detailed subsequently, the most meaningful fracture outcome measure is the proportion of patients who experience at least one clinically important fracture in an adjacent vertebra. However, this outcome is not well reported. Only five studies (Buchbinder, Farrokhi, FREE, Rousing, VERTOS) reported the overall incidence of new clinical vertebral fractures, and one of these (VERTOS) did so only for the PVP group. Blasco et al stated that 71% of the radiographic fractures in the PVP group were clinical, compared with 9% in the control group

(OR 25.67, 95% CI 3.04 to 216.8, p=0.029);[146] however, the number of patients who suffered clinical vertebral fractures was not reported. None of the other three studies which reported this outcome in both treatment groups identified a statistically significant difference between treatment groups (see Table 15).

Table 15: Incidence of clinical vertebral fractures

Study	Length of follow-up	Time from estimated fracture onset to intervention (weeks)		No of patients with incident VCF			RR (95% CI)	P value
		Intervention	Control	PVP	BKP	Control		
Buchbinder 2009[101]	6 months	Median: 9.0 (3.8-13.0)	Median: 9.5 (3.0-17.0)	3/38 (7.9%)		4/40 (10.0%)	*0.79 (0.19 to 3.30)*	*0.75*
Farrokhi 2011[147]	24 months	Median: 27 (4-50)	Median: 30 (6-54)	1/38 (2.6%)		6/39 (15.4%)	*0.17 (0.02 to 1.35)*	*0.09*
FREE[151,196]	12 months	Mean: 5.6 (4.4)	Mean: 6.4 (5.2)		21 (14%)	NR	Not calculable	
	24 months				31/149 (20.8%)	27/151 (17.9%)	*1.16 (0.73 to 1.85)*	*0.52*
Rousing 2009[139]	12 months	Mean: *1.2 (0.5-1.9)*	Mean: *1.0 (0.3-1.6)*	0/26		3/24 (12.5%)	*0.13 (0.01 to 2.44)*	*0.17*
VERTOS[153]	2 weeks	Mean: *4.2 (2.4)*	Mean: *3.8 (2.3)*	2/18 (11.1 %)		NR	Not calculable	

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

Liu reported that adjacent segment fractures occurred at 41 and 50 days after surgery in two patients in the BKP group.[148] As these fractures were reported as adverse events, they were presumably clinical rather than radiographic fractures. No such fractures were reported in the PVP group. However, it is not clear whether fractures occurred, but were not reported, in non-adjacent vertebrae. In the FREE study, at 24 months 11 patients in the BKP group (7.4%) were said to have had clinical fractures which were considered "possibly or probably related" to the intervention.[196]

Other adverse events

The included studies varied considerably in their reporting of other adverse events. Six studies (Blasco, INVEST, Liu, Rousing, VERTOS, VERTOS II) did not report any other adverse events. Farrokhi et al stated only that no emboli occurred: these were clearly envisaged as different from cement leakages, which were reported separately.

Buchbinder et al reported a number of adverse events during the first 6 months of follow-up (see Table 16). The figures appear to refer to the number of events, not the number of patients suffering the event.

Table 16: Adverse events reported from the study by Buchbinder et al[101]

Event	PVP					Control				
	1 wk	1 mo	3 mo	6 mo	Total	1 wk	1 mo	3 mo	6 mo	Total
Incident non-vertebral fracture										
Hip	0	0	1	0	**1**	0	0	0	0	**0**
Rib	1	0	1	0	**2**	2	0	0	2	**4**
Pelvis	0	0	0	0	**0**	0	1	0	0	**1**
Osteomyelitis	0	1	0	0	**1**	0	0	0	0	**0**
Tightness in back or ribcage	0	1	0	0	**1**	0	0	2	0	**2**
Pain or burning in thigh or leg	3	0	1	0	**4**	1	0	1	0	**2**
Stomach pain	1	0	0	1	**2**	0	0	1	0	**1**
Increased pain or muscle cramping around puncture site	1	0	1	0	**2**	0	0	0	1	**1**
Chest pain	3	0	0	0	**3**	0	0	0	0	**0**

The FREE study provided extensive data relating to adverse events. Data relating to serious adverse events (defined as adverse events which resulted in death, life-threatening injury, or permanent impairment, or which required extended hospital stay or intervention to prevent impairment) are summarised in Table 17. Few of these serious adverse events were considered to be related to BKP. However, a haematoma which occurred at the surgical site within 2 days of the intervention was considered to be procedure-related, as was the exacerbation of a recurrent urinary tract infection (UTI) by catheterisation, also within 2 days of surgery. The patient with the UTI also developed spondylitis near the cement in the vertebral body 376 days after surgery and was treated with antibiotics; however, the inflammation had not resolved by 24 months. None of the AEs which resulted in death (12 in

the BKP group and 11 in the control group) were considered to be related to the device or procedure.[196]

Table 17: The FREE study: patients with serious adverse events, data to 24 months[196]

	BKP	Control
Any serious AEs within 24 months	74	73
Anaemia	3	2
Back pain	5	12
Spondylitis	1‡	0
Cardiovascular and vascular disorders:		
Angina pectoris		
Arrhythmia	2	5
MI	2	3
Pulmonary embolism	5†	3
Stroke	4	1
Haematoma		
Other	1††	1
Infections:		
Clostridium infection		
Sepsis/septic shock	1	3
UTI	2‡	3
Neoplasms/cancer	7	9
Psychiatric disorders – depression	3	1
Respiratory disorders:		
Pneumonia	8	6
Dyspnoea	1	4

† 1 MI preceded surgery and resulted in death; †† Deemed to be related to BKP kyphoplasty procedure;
‡ 1 UTI was considered procedure-related: a recurrent UTI was exacerbated by catheterisation within 2 days of surgery. Spondylitis developed in the same patient near the cement in the vertebral body 376 days after surgery and was considered possibly cement-related; it was treated with antibiotics, but the inflammation had not resolved by 24 months.

Sub-groups

The evidence relating to the sub-groups specified in the scope is discussed in turn below.

Time from fracture to intervention

Of the included studies, only INVEST reported data by baseline pain duration. A post-hoc subgroup analysis of the effect of treatment on pain at 1 month by baseline pain-duration

categories found no significant difference (p comparing all 3 categories = 0.58, see Table 18).[102]

Table 18: Effect of treatment on pain at 1 month in the INVEST study, by duration of pain at baseline (data from Kallmes et al 2009[102])

Duration of pain at baseline	T1	T2	Treatment effect (95% CI)	P value
<13 weeks	NR	NR	0.8 (-0.8 to 2.4)	0.31
14-26 weeks	NR	NR	1.3 (-0.8 to 3.4)	0.23
27-52 weeks	NR	NR	0.0 (-1.7 to 1.6)	0.96

As the INVEST study was underpowered for this analysis, Staples et al undertook a meta-analysis of IPD from the INVEST and Buchbinder studies to assess the effectiveness of PVP in patients with fracture pain of recent onset (≤6 weeks) compared with pain of longer duration.[109] Because the INVEST study allowed crossover after one month, outcomes were only compared up to that time point. No statistically significant differences in RDQ scores, EQ-5D scores, or pain scores were identified between participants whose pain was of recent onset, and those whose pain duration exceeded 6 weeks (see Tables 19 to 21).

Table 19: Change from baseline in mean (SD) RDQ scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Duration of pain	PVP	Control	Adjusted† mean between-group difference (95% CI) (negative values favour intervention)	P value
≤6 weeks	-3.8 (5.9)	-4.4 (5.4)	0.2 (-3.0 to 3.4)	NS
>6 weeks	-4.2 (6.0)	-3.7 (6.3)	-1.0 (-3.0 to 1.0)	NS
All patients	-4.1 (5.9)	-3.9 (6.1)	-0.8 (-0.9 to 2.4)	NS

† Adjusted for study centre

Table 20: Change from baseline in mean (SD) EQ-5D scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Duration of pain	PVP	Control	Adjusted† mean between-group difference (95% CI) (positive values favour intervention)	P value
≤6 weeks	0.15 (0.24)	0.15 (0.30)	0.03 (-0.06 to 0.13)	NS
>6 weeks	0.11 (0.18)	0.09 (0.20)	0.03 (-0.03 to 0.09)	NS
All patients	0.12 (0.19)	0.11 (0.23)	0.03 (-0.02 to 0.08)	NS

† Adjusted for study centre

Table 21: Change from baseline in mean (SD) overall pain scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by pain duration; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Duration of pain	PVP	Control	Adjusted† mean between-group difference (95% CI) (negative values favour intervention)	P value
≤6 weeks	-3.1 (3.3)	-2.8 (4.0)	-0.1 (-1.6 to 1.4)	NS
>6 weeks	-2.7 (2.9)	-2.0 (2.7)	-0.8 (-1.8 to 0.1)	NS
All patients	-2.8 (3.0)	-2.2 (3.2)	-0.6 (-1.4 to 0.2)	NS

† Adjusted for study centre

Presence of fracture-related deformity before treatment

No data were identified relating to sub-groups with and without fracture-related deformity before treatment.

Receipt of inpatient care before treatment

None of the studies provided information on the number of patients who were inpatients at the time of randomisation, and no data were identified relating to this sub-group.

Baseline pain severity

In the absence of data relating specifically to patients who had received inpatient care immediately preceding the intervention, it may be relevant to note that Staples et al's analyses of IPD from the Buchbinder and INVEST studies include patients grouped by baseline pain severity. While p values were not reported, no statistically significant differences in RDQ

scores, EQ-5D scores, or pain scores, were identified between participants with severe pain (score ≥8 on a 0-10 rating scale) or mild to moderate pain (score <8) at baseline.[109] In both treatment groups, the decrease in pain was greater in the sub-group which had more severe pain at baseline than in the sub-group with less severe baseline pain (see Tables 22 to 24), but this presumably simply reflects a greater potential for improvement.

Table 22: Change from baseline in mean (SD) RDQ scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Baseline pain score	PVP	Control	Adjusted† mean between-group difference (95% CI) (negative values favour intervention)	P value
<8	-4.2 (6.0)	-4.4 (6.4)	-0.2 (-2.5 to 2.1)	NS
≥8	-4.1 (5.9)	-3.3 (5.6)	-1.4 (-3.9 to 1.2)	NS
All patients	-4.1 (5.9)	-3.9 (6.1)	-0.8 (-0.9 to 2.4)	NS

† Adjusted for study centre

Table 23: Change from baseline in mean (SD) EQ-5D scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Baseline pain score	PVP	Control	Adjusted† mean between-group difference (95% CI) (positive values favour intervention)	P value
<8	0.09 (0.17)	0.07 (0.21)	0.02 (-0.04 to 0.09)	NS
≥8	0.16 (0.21)	0.15 (0.25)	0.05 (-0.03 to 0.12)	NS
All patients	0.12 (0.19)	0.11 (0.23)	0.03 (-0.02 to 0.08)	NS

† Adjusted for study centre

Table 24: Change from baseline in mean (SD) overall pain scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures, by baseline pain severity; meta-analysis of data from the Buchbinder and INVEST studies (data from Staples et al 2011[109])

Baseline pain score	PVP	Control	Adjustedi† mean between-group difference (95% CI) (negative values favour intervention)	P value
<8	-1.9 (2.8)	-1.1 (2.8)	-0.8 (-1.9 to 0.3)	NS
≥8	-3.9 (2.9)	-3.5 (3.2)	-0.3 (-1.5 to 0.8)	NS
All patients	-2.8 (3.0)	-2.2 (3.2)	-0.6 (-1.4 to 0.2)	NS

† Adjusted for study centre

The evidence relating to pain severity prior to PVP therefore suggests that there is no reason to suppose that outcomes would differ in patients who were inpatients prior to treatment and those who were not. This view is strengthened by the fact that receipt of inpatient care following VCF may be influenced by factors other than clinical factors such as pain severity: patients who are bedridden with severe pain may not be hospitalised if they have adequate support networks in terms of both family/friends and community services. No sub-group data are available for BKP.

Summary of evidence for the clinical effectiveness of PVP and BKP

The volume of available evidence of clinical effectiveness is greater for PVP than for BKP, and the methodological quality of some of that evidence is also higher than that of any study of BKP. Thus, the studies at least risk of bias are the double-blinded Buchbinder and INVEST studies comparing PVP with an OPLA. The studies which compare PVP with conservative management (Blasco, Farrokhi, Rousing, VERTOS, VERTOS II) vary in quality, that by Farrokhi et al being at least risk of bias.

The FREE study, the only study to compare BKP with conservative management, is at risk of bias because of the lack of blinding of patients and outcome assessors, the relatively high loss to follow-up, the unexpected imbalance in drop-outs, and selective reporting of outcomes.

The study by Liu et al, the only study to compare PVP with BKP, is poorly reported and potentially at risk of bias from a number of sources. It is also underpowered to identify statistically significant differences in effectiveness between the two interventions.

In relation to PVP, the studies least at risk of bias (Buchbinder, INVEST) found no significant differences between treatment groups in terms of change from baseline in health-related quality of life other than in terms of the total QUALEFFO score at 1 week in the Buchbinder study; this favoured PVP. No significant differences were observed in any measure of functional status or pain (whether measured in terms of mean pain scores or numbers of patients reporting clinically meaningful improvements in pain). Although the INVEST study reported a trend towards a greater number of patients in the PVP group reporting a clinically meaningful improvement in pain at one month, pooled data from the Buchbinder and INVEST studies indicate that, after adjusting for baseline opioid use, patients randomised to PVP were more likely than those randomised to the OPLA to be taking opioids at 1 month. Consequently, it is impossible to exclude the possibility that PVP was associated with worse outcomes which were masked by greater opioid use.

What evidence there is from the open-label studies of PVP (Farrokhi, Rousing, VERTOS, VERTOS II) regarding HRQoL is not consistent: VERTOS and VERTOS II suggest that PVP is associated with better HRQoL, as measured by the QUALEFFO, whereas Blasco found no significant difference between treatment groups. The data reported by Rousing et al indicate that conservative management is generally associated with better HRQoL, as measured by the EQ-5D and DPQ. By contrast, the evidence from these studies relating to functional status appears to favour PVP: the most convincing evidence comes from Farrokhi, the unblinded study at least risk of bias, which found that, as measured by a modified version of the ODI, PVP was associated with significantly improved functional status at all times from 1 week to 36 months. In the Farrokhi study, mobility at 1 day was also dramatically better in the PVP group than in the control group. The Blasco, Farrokhi, Rousing, and VERTOS II studies found that PVP was associated with significant improvements in pain, although in the study by Blasco et al statistical significance was only seen at 2 months; moreover, in those studies which report analgesic use (Blasco, VERTOS II), these improvements do not appear to be associated with increased analgesic use in the PVP group. Farrokhi also found that PVP was associated with sustained improvements in vertebral body height and angular deformity; however, Blasco found no significant difference between groups in VBH.

The unblinded FREE study of BKP found that, compared with conservative management, BKP was associated with significantly greater improvements from baseline in HRQoL, although these diminished over time. It was also associated with an improvement in functional status as measured by the RDQ at one and 12 months, but not at 24 months, and with a significantly reduced risk of needing walking aids or bed rest/restricted activity at one month, but not at 12 months. BKP was also associated with significant short- and medium-

term reductions in pain, and with significant reductions in opioid use up to, but not beyond, 6 months. The effect of BKP on VBH was not reported; a statistically significant improvement in kyphotic angle was reported but its clinical significance is not clear.

In theory, the additional benefits of BKP compared with PVP are:

- The restoration of vertebral height and spinal alignment
- A lower incidence of cement leaks because the cement is injected at lower pressure.

In the study by Liu et al, BKP was said to be associated with greater improvements than PVP in VBH and angular deformity (both, $p<0.001$). Cement leaks were not reported. While data from the included RCTs and observational studies do indeed suggest that the incidence of cement leaks is lower with BKP than with PVP, because this finding is not derived from a randomised head-to-head comparison, it is possible that it may reflect differences in patient selection.

The study by Liu et al did not attempt to assess HRQoL or functional status. It did not identify a statistically significant difference between PVP and BKP in terms of pain, nor was it powered to do so.

Subgroup analyses conducted by Staples et al.[109] using individual patient data from the Buchbinder and INVEST studies found no differential benefit for PVP in relation to either baseline pain duration or pain severity. No subgroup data relating to BKP are available.

Adverse effects and contraindications: observational studies

As noted in section 5, since RCTs generally perform poorly at detecting long-term or rare AEs, it was decided to examine large case series ($N \geq 200$) and individual case reports in order to gain a rough estimate of incidence of more common AEs from large cohorts, whilst also scoping the rarer but serious events which are often published as individual case reports. We hoped in this way to be able both to identify the range of potential AEs associated with PVP and BKP, and also quantify the incidence of the more common AEs. Previous systematic reviews of adverse events have been criticised for focusing on predefined adverse events – an approach which may miss unexpected, but potentially important information.[177] By our inclusion of case reports relating to AEs which were not reported in the large case series, combined with our decision not to define adverse events of interest *a priori*, we sought to avoid this pitfall.

Our searches identified no publications of registry data which were specific to patients with osteoporotic VCF. If such studies had been identified, they would have been included either

as large case series, if they presented data relating to all patients undergoing PVP or BKP, regardless of outcome, or as an agglomeration of individual case reports if they only presented data relating to patients who had suffered AEs. However, the Medtronic submission[34] included two unpublished reports which compared mortality and complication risks for operated and non-operated patients with osteoporotic VCF; their findings are summarised below.

Registry data

The Medtronic submission included claims-based data relating to the US Medicare population for the years 2005 to 2009 inclusive[174] and to subscribers to a major German health insurance fund (AOK Niedersachsen) for the years 2005 to 2010 inclusive.[175,175] [redacted]

[redacted]

[redacted]Data from these studies are summarised in Appendix 10; they will also be discussed below in the relevant contexts. [redacted]

[redacted]Additionally a formal critique of the data is providedin Appendix 12, and summarised later.

Large case series

The electronic searches revealed 14 large case series (N≥200) which reported AEs associated with vertebroplasty and balloon kyphoplasty. Five of these articles included data from procedures performed on VCFs of non-osteoporotic origin. In these cases, data from osteoporotic VCFs only was requested from the corresponding authors. One corresponding author (Mpotsaris) replied to the request, giving a total of 10 case series which provided useable data. The AE data for balloon kyphoplasty and vertebroplasty are summarised in Table 25 and Table 26 respectively.

Table 25. Adverse events in large (n≥200) balloon kyphoplasty case series

Study	Total patients (n)	Mean follow-up duration (range)	Total treated VB	Total n treated levels with leakage	In patients (n)	Location(s) of leakages	Neurological complications	Other
Blattert et al. 2010[205]	314	Minimum 2 years	352	32	NR	Epidural space (n=6), others NR	NR	Intraoperative balloon perforation (n=6)
Diel et al. 2010[206]	320	7 months (20 to 389 days)	391	70	62	Intervertebral disc (n= 28); paravertebral vessels (n=13); epidural space (n=4); other (n=25)	Radiculopathy due to cement extrusion (n=3)	Intraoperative balloon rupture (n=1); fracture of vertebral wall with displacement of balloon catheter (n=1); interruption of surgery for unspecified reason (n=1); nonspecified complication (n=1)
Majd et al. 2005[207]	222	21 months (6 months to 36	360	38	NR	NR	Radiculopathy caused by cement	Unspecified medical complications

Study	Total patients (n)	Mean follow-up duration (range)	Total treated VB	Total n treated levels with leakage	In patients (n)	Location(s) of leakages	Neurological complications	Other
		months)					leakage into foramen (n=1)	(n=10): “most . . . were related to pre-existing cardiac, pulmonary, or liver disease” 1 patient required surgical debridement, irrigation, and closure of wound 3 weeks post-procedure. 1 infection at kyphoplasty site 2 months post-procedure, leading to abscess formation at

Study	Total patients (n)	Mean follow-up duration (range)	Total treated VB	Total n treated levels with leakage	In patients (n)	Location(s) of leakages	Neurological complications	Other
								L3, and consequent cardiovascular failure and death.

VB, vertebrae; VCFs, vertebral compression fractures; NR, not reported

Table 26. Adverse events in large (n≥200) vertebroplasty case series

Study	Total patients (n)	Mean follow-up duration (range)	Total treated VB	Total n treated levels with leakage	In patients (n)	Location(s) of leakages	Neurological complications (in patients, n)	Other
Álvarez et al. 2005[208]	260	12 months (3 weeks to 96 months)	423	305	NR	Spinal canal (n=3); vertebral disc (n=43); lumbar venous plexus (n=44); epidural veins (n=132)	Transitory radicular pain (n=12); transitory parapesia (n=1)	Rib fractures in 5 patients
Diel et al. 2009[209]	203	2 months	1137	126	NR	NR	NR	Temporary hypotension (n=8); pulmonary embolism (n=1)
Evans et al. 2003[210]	245	Median 7.2 months (IQR: 3.1 months to 13.6 months)		NR	NR	NR	Transitory radicular pain (n=2)	Rib fractures in 7 patients; post-procedural worsening of pain (n=3)

Study	Total patients (n)	Mean follow-up duration (range)	Total treated VB	Total n treated levels with leakage	In patients (n)	Location(s) of leakages	Neurological complications (in patients, n)	Other
Lee & Chen, 2004[211]	200	NR	200	29	29	Disc space or paravertebral space (distribution NR)	NR	NR
Masala et al. 2009a[212]	285	Up to 3 years	429	21	NR	Disc space or paravertebral veins (distribution NR)	NR	NR
Mpotsaris et al. 2011[a213]	896	Up to 12 months	NR	NR	108	Paravertebral venous plexus	NR	NR
Ryu & Park, 2009[214]	215	15 months (6 months to 22 months)	383	NR	187	Epidural space (n=157); paravertebral space (n=18); intradiscal space (n=12)	NR	NR

[a] case series included VCFs of non-osteoporotic origin (e.g. traumatic, metastases of malignant tumours), data shown relates specifically to patients with osteoporosis

VB, vertebrae; VCFs, vertebral compression fractures; NR, not reported

Serious adverse events related to balloon kyphoplasty and vertebroplasty, though relatively rare, are of sufficient importance to warrant consideration in clinical decision making. Statistical aggregation of data relating to the more common AEs was not possible because, as shown in Table 25 and Table 26, the data were heterogeneous in terms of what was reported, and how it was reported.

All-cause mortality

No deaths were noted in the large observational studies of PVP. However, one procedure-related death was noted by Majd and colleagues[207] in a case series of 222 patients who had 360 vertebral bodies treated by BKP. This patient developed an infected shunt and subsequent abscess formation at the site of kyphoplasty. He underwent a discectomy with anterior plus posterior spinal fusion and instrumentation, but did not recover well and subsequently died from cardiovascular failure. It is noteworthy that this patient had previously received a kidney transplant and was taking anti-rejection medications and prednisone.

[redacted] (for details, see Appendix 10).

A formal critique of the evidence on mortality provided by Medtronic is provided in Appendix 12, with a summary presented here. Observational data can be subject to confounding factors, although methods to adjust for these such as regression analyses using observed variables as covariates and propensity matching exist. However, neither method can produce a robust estimate of the variable of interest if there is selection of the intervention provided based on unobserved data. Where this may be likely, instrumental variable methods using a variable correlated with an intervention, but which is only correlated with the outcome through its effect on the intervention can be employed. However, the validity of an instrumental variable is subjective and can be open to debate.

Evidence on mortality benefit associated with BKP and PVP was submitted by Medtronic in the form of four studies, all using observational data, two from a claims database from the US and two from a health insurance fund in Germany. A variety of methods are used, including Cox regression using covariates, matching methods and IV estimation. The results involved paired comparisons between different groups rather than simultaneous comparisons of the three treatments, which may introduce inaccuracy. It is unclear how generalisable these results are to patients treated in England and Wales.

[redacted]

[redacted]

[redacted]

[redacted]

In summary, it is possible that there is a causal difference in mortality between patients treated using OPM and patients receiving BKP or PVP given the size of the effect. Appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. It is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment based on the data presented in the studies

included here. There is also considerable uncertainty were BKP and PVP assumed to have a mortality benefit, in whether OPLA would also produce a mortality benefit but no data are available on this

Table 27. Summary of results estimating a mortality benefit associated with BKP or PVP

Edidin et al (2011[215]) Mortality risk 4 years	Group	Comparison	Cox regression Adjusted HR (95% CI)	IV at 3 years Relative increase In survival	
	All	OP vs. OPM	0.63 (0.62-0.64)		
		BKP vs. OPM	0.56 (0.55-0.57)		
		PVP vs. OPM	0.76 (0.75-0.77)		
		BKP vs. PVP	0.77 (0.75-0.78)		
	Survival>1 year	OP vs. OPM	0.82 (0.81-0.84)		
		BKP vs. OPM	0.76 (0.74-0.77)		
		PVP vs. OPM	0.93 (0.91-0.95)		
		BKP vs. PVP	0.82(0.80-0.85)		
	Operated	BKP vs. PVP		11.82%	
Exponent (2012)[174] Mortality risk 5 years	Group	Comparison	Cox regression Adjusted HR (95% CI)	Propensity score Matching and Cox regression HR (95% CI)	
	All	OPM vs. OP[(1)]	[redacted]	[redacted]	
		OPM vs. BKP			
		OPM vs. PVP			
		BKP vs. PVP			
	OVCF	OPM vs. OP[(1)]			
		OPM vs. BKP			
		OPM vs. PVP			
		BKP vs. PVP			
	OVCF Survival>1 year	OPM vs. OP[(1)]			
		OPM vs. BKP			
		OPM vs. PVP			
		BKP vs. PVP			
Lange and Braun (2012a,b[175,216]) Mortality risk 5 years	Group	Comparison	Cox regression Adjusted HR (95% CI)	Propensity score Matching and Cox regression HR (95% CI)	Propensity score Matching Difference in survival rates % [p-value]
	OVCF	OP vs. OPM	[redacted]		
		BKP vs. PVP		[redacted]	[redacted]
	OVCF Survival>1 year	OP vs. OPM			
		BKP vs. PVP			

[(1)]Results reported in the appendix but not reported in the main text.

Note that HRs have been reported rather than statistics such as median or mean survival. This is due to relatively large numbers of patients remaining alive at the end of the follow-up

period. For example in the Edidin et al[215] publication median survival had not been reached at four years since VCF diagnosis in any of the arms.

Symptomatic and asymptomatic cement leakage

The most common risk associated with vertebral augmentation procedures is cement leakage outside the target vertebral body. The majority of articles reported the incidence of cement leakage in terms of treated vertebral bodies, while four reported it in terms of treated patients. Only Diel's study of kyphoplasty[206] and Lee's study of vertebroplasty[211] provided leakage incidence data for both treated vertebrae and treated patients. Taken in isolation, data relating to either the number of vertebrae or the number of patients are potentially misleading and could introduce systematic bias towards underreporting of incidence.

The location of cement leakages has important implications for safety: intradiscal leakages are unlikely to lead to morbidity, but leakages into the epidural space or venous system have the potential to cause serious complications.[217] Three studies (Majd,[207] Diel,[209] Evans[210]) did not report the location of cement leakages. Poor reporting of follow-up duration and completeness was also a problem for interpreting these data. Lee[211] did not report follow-up duration while, in most of the other studies, it was unclear what proportion of the cohort was lost to follow-up at what time points, and why.

When reported in terms of treated vertebral bodies, the incidence of leakage ranged from 5%[212] to 72%[208] for vertebroplasty, and from 9%[205] to 18%[206] for kyphoplasty. By contrast, when reported in terms of treated patients, leakage incidence was higher, ranging from 12%[213] to 87%[214] for vertebroplasty. Only one kyphoplasty study[206] reported incidence in terms of treated patients: it reported a rate of 19%. It is not clear why such wide variations in incidence were observed, but factors such as practitioner skills and experience, clinical setting, cement viscosity, and thoroughness of follow-up may have played a part.

Epidural leaks appeared to be common in vertebroplasty cohorts. Ryu and Park[214] reported epidural leaks in 157 of 215 treated patients (73%), and Álvarez et al.[208] reported three leaks into the spinal canal and 132 into the epidural veins, in a cohort of 260 patients with 423 treated vertebrae (52% of patients). These complications did not appear to be as common in kyphoplasty cohorts. Blattert et al.[205] and Diel et al.[206] reported six and four leaks into the epidural space in cohorts of 314 and 320 respectively (2% and 1% respectively). Because of the nature of follow-up in the cohorts, the long-term clinical implications of these cement leaks are unknown. Several investigators undertook long-term follow-up: Masala and colleagues[212] reported a follow-up duration of up to 3 years, although only 68 patients (24%)

had data available at that time point; Majd et al.[207] reported a follow-up duration of up to 36 months, although data were not available on how many patients had data available at given time-points; Blattert and Josten[205] reported a minimum 2 year follow-up, and did not report any missing data; and Álvarez et al.[208] reported follow-up of up to 96 months though, again, it was unclear how many patients were followed up at particular time points.

Other reported AEs which may be related to cement leakage included pulmonary embolism,[209,218] radiculopathy[206,207] (which in the study by Diel[206] was specifically due to cement extrusion), temporary radicular pain,[210] and temporary and permanent motor deficits or parapesia of the legs.[208]

Intraoperative balloon rupture

Intraoperative balloon rupture appears to be a relatively rare complication of BKP: in the two studies which report it (Blattert et al.[205] and Diel et al.[206]), it occurred in 6 out of 352 procedures (1.7%) in Blattert et al.'s cohort,[205] and in 1 out of 391 (0.3%) in Diel et al's study.[206] Neither of these studies discussed the clinical implications of balloon rupture. However, Saliou et al.[219] discussed some of the potential implications in a smaller case series in which it was more common (n=51, treated levels: 75, balloon rupture in 5 vertebrae of 5 patients). Although no symptomatic complications due to balloon rupture were observed in that study, the authors point out that this complication could lead to contrast leakage, procedural delay, or gas embolism.

Other peri- and post-operative complications (including infection)

In the included case series, peri- and post-operative complications were relatively rare. Majd et al[207] reported 10 medical and three surgical complications in 222 patients undergoing BKP. Most of the medical complications related to pre-existing cardiac, pulmonary, or liver disease. In one case, a patient was treated with local anaesthesia because medical comorbidities made general anaesthetic inadvisable, and developed electrocardiogram abnormalities during the procedure; treatment of a second VCF had to be postponed for four days, while the patient was assessed by a cardiologist. In addition to one case of infection discussed under mortality above, and one cement leak causing radiculopathy, also discussed above, the surgical complications included one patient who needed surgical debridement, irrigation, and closure of the wound 3 weeks after BKP.[207]

Diel et al reported one instance where the vertebral wall was fractured, with displacement of the balloon catheter, in a patient undergoing BKP,[206] and 8 cases of temporary hypotension

following cement injection in 202 patients undergoing PVP for osteoporotic fracture (3.9%).[209]

Incidence of new vertebral fractures

New vertebral fractures have been identified as an important source of postoperative morbidity among people with osteoporotic VCFs treated with PVP or BKP. The observational studies by Harrop,[220] Kulcsar,[221]Tseng,[222] and Uppin[223] specifically set out to study the overall incidence of new vertebral fractures in osteoporotic patients following PVP or BKP. However, as these patients are by definition at increased risk of vertebral fracture, the data are difficult to interpret in the absence of a control group of similar patients who have not undergone PVP or BKP. Similarly, although a number of retrospective reviews of new vertebral fractures in patients treated with PVP or BKP were identified,[164,220-236] and reported incidence rates ranging between 6.8% over a 25.6 month follow-up period[229] to 22.2% during a 1-year follow-up,[233] it is difficult to know how to interpret these data. However, it should be noted that, because new fractures were generally identified only when patients returned to clinic with recurring back pain, the reported figures probably represent a conservative estimate of true fracture incidence.

Arguably of greater relevance was the finding from the case series by Harrop,[220] Kulcsar,[221]Tseng,[222] and Uppin[223] that new VCFs are significantly more likely to occur in vertebrae adjacent to treated levels than in nonadjacent vertebrae. Although two reports found a similar crude incidence rate of adjacent and nonadjacent fractures,[232,234] it should be added that patients would typically have a greater number of nonadjacent vertebral bodies that could fracture, so even these data may represent a greater likelihood of fracture at adjacent levels.[236] In addition, some studies show that, following vertebral augmentation, adjacent fractures are likely to occur sooner than nonadjacent fractures. Donovan et al.[224] reported the case of a 50-year-old woman who developed several new fractures eight days after a kyphoplasty procedure, and concluded that the "temporal relationship between the kyphoplasty procedure... with documented fractures of six adjacent vertebrae... is highly suggestive of causality" (p. 712). A larger retrospective analysis of time between vertebroplasty and new adjacent fractures[234] found times to diagnosis of new adjacent and nonadjacent fractures of 55 and 127 days respectively ($p<0.0001$). Further evidence was supplied by Mudano and colleagues,[237] who compared the rate of new fractures in a cohort of patients treated with PVP or BKP, against a cohort of patients with VCFs and no cement augmentation. A significantly higher incidence was observed in the treated cohort at 90 days (adjusted odds ratio (OR): 6.8; 95% CI: 1.7–26.9) and 360 days (adjusted OR: 2.9; 95% CI: 1.1–7.9).

A number of prognostic factors have been associated with higher risk of subsequent vertebral fractures: these include increased age and number of treated vertebrae,[228] presence of clefts in the treated VCFs,[235,238] and spinal instability measures.[227] There is also a growing body of evidence suggesting biomechanical explanations for the higher rate of adjacent fractures. A number of studies have demonstrated that bone cement can increase the stiffness of the treated vertebra, resulting in an increase in loading on the adjacent vertebrae.[239-242] Cement leakage may also play a part: Han et al.[225] found that, when adjacent VCFs occurred, fractures were more likely to be close to extraneous cement. In contrast to these studies, Farooq et al.[243] demonstrated that vertebroplasty could partially reverse fracture-induced changes including decompression of the adjacent nucleus, and higher neural arch load-bearing.

However, in the absence of well-controlled randomised studies, neither time from surgery to new VCF, nor a higher incidence of adjacent vertebral fractures (compared with nonadjacent fractures), can be considered as definitive evidence of causation.

Rib fractures

Two large observational studies reported rib fractures related to vertebroplasty: Álvarez et al.[208] reported five fractures in a cohort of 260 patients, while Evans et al.[210] reported seven in a cohort of 245. No rib fractures were reported in the kyphoplasty case series.

Refracture of treated vertebrae

It has been suggested that a treated vertebra may refracture either because too little cement was injected, or because the vertebra was extremely fragile and therefore at risk of refracture even when adequate quantities of cement were injected.[166] However, as none of the included case series reported this complication, incidence is likely to be low.

PVP and BKP may be associated with transitory increase in post procedural pain. However, among the large case series, only Evans et al[210] reported this complication: 3 patients from a cohort of 245 experienced worsening of pain, although no biomechanical causes could be found. While the other case series did not report worsening of pain as an AE, it was unclear whether this was because it did not occur, or because the authors did not view transitory increases in pain as an AE per se.

Need for repeat procedure

A small proportion of patients may require repeat vertebroplasty because of adverse events. Yang et al[244] presented data relating to 22 patients who required repeat PVP: 20 out of 1523 consecutive patients who underwent VP for osteoporotic fracture in their centre between

2000-2006 (1.3%), who had recurrent back pain after a short period of pain relief following first-time vertebroplasty, and a further two patients with neurological deficits following first-time vertebroplasty who were referred from other hospitals for revision surgery. The reasons for revision surgery, and the nature of the intervention required, are presented in Table 28. Most patients were discharged from hospital within two weeks, but those with infections required longer hospitalisation because they received at least a six-week course of parenteral antibiotics. Four patients required a third surgical procedure.

Table 28: Complications requiring revision surgery following vertebroplasty (data from Yang et al[244])

Complication	Number of patients	Intervention required
Residual vacuum cleft or poor cement augmentation	5	Repeat vertebroplasty
Poor cement augmentation and progressive kyphosis and instability	2	Posterior surgery (instrumentation and fusion)
Infection (pyogenic spondylitis)	8	Anterior and posterior surgery
Cement dislodgment	3	Anterior or anterior and posterior surgery
Cement fragmentation	2	Anterior surgery
Neurological deficit	2	Anterior and posterior surgery

Case reports

In general, case reports were included only if they reported adverse events which had not been reported in the larger observational studies. However, an exception was made in the case of pulmonary cement embolism, the most commonly reported complication of vertebral augmentation, where all identified case reports were included in order to indicate subsequent therapy, if any.

Pulmonary cement embolism

The search identified 46 case reports and 47 patients in whom a pulmonary cement embolism caused by venous PMMA leakage was detected. These reports related to 41 vertebroplasty and five kyphoplasty procedures. Four deaths due to pulmonary embolism were reported in vertebroplasty patients; no deaths were identified in kyphoplasty patients. Sixteen embolisms were reported as asymptomatic, while 29 were symptomatic; for the remaining two, no details were provided on symptomatology. Symptomatic manifestations of pulmonary embolism

include dyspnoea, tachycardia, chest pain, dizziness and sweating. Asymptomatic pulmonary embolism is more difficult to detect and, furthermore, it is difficult to gain understanding the long-term clinical implications of these silent pulmonary emboli from the available data. The case study data relating to pulmonary embolism are summarised in Table 29.

Table 29. Case reports of pulmonary embolism after vertebral augmentation

Authors	**Number of patients**	**Procedure (PVP/BKP)**	**Symptomatic/ asymptomatic**	**Therapy**
Abdul-Jalil[245]	2	PVP	1 symptomatic, 1 asymptomatic	Low dose heparin
Agko[246]	1	BKP	Asymptomatic	Surgical embolectomy
Baumann[247]	1	PVP	Asymptomatic	Coumarin 3 months
Bernhard[248]	1	PVP	Asymptomatic	-
Biega[249]	1	PVP	Asymptomatic	-
Bonardel[250]	1	PVP	Asymptomatic	Coumarin 6 months
Cadeddu[251]	1	PVP	Asymptomatic	-
Caynak[252]	1	PVP	Symptomatic	Anticoagulants and pulmonary physiotherapy
Chen[a253]	1	PVP	Symptomatic	CPR
Dastidar[254]	1	PVP	Symptomatic	Inferior vena cava filter placement

Authors	**Number of patients**	**Procedure (PVP/BKP)**	**Symptomatic/ asymptomatic**	**Therapy**
Finch[255]	1	PVP	Symptomatic	-
Francois[256]	1	PVP	Symptomatic	Coumarin 6 months
Freitag[257]	1	PVP	Asymptomatic	Coumarin 6 months
Grahe[258]	1	PVP	Symptomatic	Anticoagulation and oxygen therapy
Harris[259]	1	PVP	Symptomatic	-
Jang[260]	2	PVP	Symptomatic	Anticoagulants and heparin
Kim[261]	1	PVP	Symptomatic	Surgical embolectomy
Kovalenko[262]	1	PVP	-	-
Lee[263]	1	PVP	Symptomatic	Surgical embolectomy
Leroux[264]	1	PVP	Symptomatic	-
Liliang[265]	1	PVP	Symptomatic	None
Lim[266]	1	PVP	Symptomatic	Surgical embolectomy
Lim[267]	1	PVP	Symptomatic	Surgical embolectomy
MacTaggert[268]	1	PVP	Asymptomatic	-
Moll[269]	1	BKP	Symptomatic	Anticoagulation 3

Authors	Number of patients	Procedure (PVP/BKP)	Symptomatic/ asymptomatic	Therapy
				months
Monticelli[a270]	1	VP	Symptomatic	CPR
Moon[271]	1	VP	-	Anticoagulation
Müller[272]	1	BKP	Asymptomatic	-
Neuwirth[273]	1	PVP	Asymptomatic	-
Perrin[274]	1	PVP	Symptomatic	Low dose heparin
Pleser[275]	1	PVP	Asymptomatic	Heparin and coumarin 6 months
Pott[276]	1	PVP	Symptomatic	Low dose heparin
Quesada[277]	1	PVP	Asymptomatic	-
Radcliff[278]	1	BKP	Symptomatic	Conservative treatment
Righini[279]	1	PVP	Symptomatic	Coumarin 6 months
Schneider[280]	1	PVP	Asymptomatic	-
Schoenes[281]	1	PVP	Symptomatic	Surgical embolectomy
Scroop[282]	1	PVP	Symptomatic	None
Seo[283]	1	PVP	Asymptomatic	Surgical embolectomy

Authors	Number of patients	Procedure (PVP/BKP)	Symptomatic/ asymptomatic	Therapy
Shalshin[284]	1	BKP	Asymptomatic	Short-term enoxaparin
Son[285]	1	PVP	Symptomatic	Surgical embolectomy
Stricker[a286]	1	PVP	Symptomatic	Definitive airway
Torres Machi[287]	1	PVP	Symptomatic	Anticoagulation
Tozzi[288]	1	PVP	Symptomatic	Coumarin 3 months
Yoo[a289]	1	PVP	Symptomatic	Surgical embolectomy
Zaccheo[290]	1	PVP	Symptomatic	Low dose heparin

PVP, vertebroplasty; BKP, kyphoplasty

[a] Death due to pulmonary embolism

Postoperative infection

A number of case reports [291-302] have described postoperative infectious complications. While such infections can occasionally be managed with a medical approach, they often necessitate further surgical intervention.[293,294,296-298,300,302] One team[292] reported the death of a patient from septic multiple organ failure after antibiotic treatment and local surgical interventions.

Other adverse events

A number of case reports noted rare, but serious cardiovascular complications related to vertebral cement augmentation, including cardiac perforation,[281,285,303,304] inferior vena cava syndrome,[305] venous air embolism,[306] vena cava thrombus,[307] acute pericarditis,[308] lumbar artery pseudoaneurism,[309] and stroke.[310]

Biafora et al.[311] reported an injury to a segmental branch of the L4 lumbar artery in an 84-year-old patient: this manifested clinically in bleeding from the kyphoplasty site, and was successfully treated with torpedo embolisation of a small branch of the right L4 lumbar artery.

Heo and Cho[312] reported an L2 segmental artery injury, which was also successfully treated with endovascular embolisation. Hard et al.[313] reported a transpedicular needle penetrating the margin of the T5 vertebral body by 15mm. Injury to the posterior aortic wall was confirmed, and an improvised injection of PMMA was used to seal the aortic wall. No related complications were seen during two years of follow-up.

Ozturk et al.[314] reported a case of irreversible complete paraplegia due to cement leakage into the spinal canal. Lee et al.[315] presented a case of complete motor and sensory deficits at T11 due to cement leakage, which was treated with surgical decompression. Birkenmaier et al.[316] reported a transitory paraplegia in an 82-year-old patient following a massive epidural haematoma compressing the cauda equine and the conus medullaris. The haematoma was drained, resulting in the loss of 3L of blood and requiring transfusion of packed red blood cells and fresh-frozen plasma. However, 48 hours post-procedurally, full neurological function had been regained. Lopes and Lopes[317] also reported paraplegia due to spinal cord and root compression which was successfully remedied with surgical decompression.

Lim et al.[318] reported two cases of subarachnoid haemorrhage: both patients were treated successfully with medical management. Other rare complications included heterotopic ossification,[319] addisonian crisis,[320] lumbar disc herniation,[321] posterior spinal epidural abscess,[322] and a fatal fat embolisation with no evidence of cement leakage.[323]

Summary and discussion of data relating to adverse events

The evidence drawn from the included RCTs, case series, and case reports suggests that PVP and BKP may be associated with a number of adverse events. Treatment-related deaths appear to be rare, but cement leakage is common, particularly with PVP: pooled data from the RCTs indicate an incidence of 44% of treated vertebrae for PVP and 27% for BKP (see Table 13), while the case series indicate a range of 5% to 72% for PVP and 9% to 18% for BKP (see Tables 25 and 26). While many cement leaks were not associated with immediate clinical complications, others were associated with serious problems such as pulmonary embolism, radiculopathy, and temporary or permanent motor deficits. A number of procedure-related deaths have been noted. Moreover, there is as yet no good evidence to prove that that leaks which are asymptomatic in the short term do not have long-term implications.

Peri- and post-operative complications other than cement leak appear to be rare, though potentially serious. In particular, infectious complications are potentially fatal, and frequently require treatment with further surgical intervention. To reduce the risk of such complications, it has been recommended that PVP or BKP should not proceed until the patient has made a

complete recovery from any existing infections, and that, in cases of recent infection, either antibiotics should be prescribed on a long-term basis to avoid deep infection, or a cement-antibiotic mixture should be used.[293] Intraoperative balloon perforation during kyphoplasty seems unlikely to lead to any serious complications. Nevertheless, Saliou et al[219] have suggested a number of methods to minimise the incidence of rupture: 1. Purge any trapped air from the balloon to prevent gas embolism; 2. Increase balloon inflation pressure very slowly, to allow the balloon to adapt to the solid, sharp bony environment, and 3. Use a curette to break bone bridges and constitutive bone fragments before inflating the balloon.

While it seems likely that PVP and BKP may be associated with increased rates of new vertebral fractures, and in particular adjacent fractures, as yet the quality of the evidence for this is not good.

It is also unclear which of PVP or BKP is the safer of the two approaches to vertebral augmentations, as direct comparisons were unavailable. However, Yang et al[324] conducted a review which found that rates of specific complications (cement leakage, new compression fractures, pulmonary embolism, and radiculopathy) were all significantly higher with vertebroplasty than with kyphoplasty (all $p<0.05$). They also found that cement leakage rates were lower in procedures carried out in neurosurgery departments (20.6%) and orthopaedic departments (24.7%) than in radiology departments (52.9%). This could, however be confounding if the vertebroplasties were carried out by the radiologists. In addition, Medtronic claim that, in BKP, the creation of a cavity within the vertebral body allows for the insertion of a pre-known volume of a more viscous cement at a lower pressure, which reduces the risk of cement leakage and consequent complications compared with PVP.[34]

None of the included studies referred to the radiation risks to patients associated with PVP and BKP. These risks, though low, are not trivial. Perisinakis et al.[325] estimated, on the basis of a case series of 11 patients undergoing kyphoplasty with fluoroscopic guidance, a rate of 741 fatal cancers and 5.4 hereditary effects per million treated patients. However, Fitousi et al.[326] found a relatively high level of radiation exposure in a case series of 11 patients undergoing vertebroplasty with fluoroscopy, and estimated a fatal cancer risk of 1 in 580 and a risk of hereditary effects of 1 in 20,000.

Finally, it should be noted that, although PVP and BKP may be associated with the adverse effects discussed above, the alternative treatment (conservative management with analgesics, back bracing and bed rest) is linked to a number of potentially serious complications. Bed rest can lead to muscle wasting and deconditioning, and these effects have been associated with

deep vein thrombosis, pulmonary emboli, reduced muscle blood flow, red cell volume, capillarization, and oxidative enzymes.[327,328] Narcotic analgesics are associated with a number of undesirable side-effects including cognitive impairment and nausea, while NSAIDs are associated with gastrointestinal problems.[54] The registry studies indicate that, [redacted] (for details, see Appendix 10). [redacted]

- Discussion of clinical effectiveness

  Internal validity

  The evidence for the clinical effectiveness of PVP is not consistent. The best-quality studies, the blinded Buchbinder and INVEST studies, show no benefit, whereas some benefit is seen in the lower-quality unblinded studies.

  The unblinded FREE study suggests some benefit from BKP, although any benefits diminished over time.

  Various suggestions have been put forward to explain the inconsistency between the results of the blinded and unblinded studies of PVP. These suggestions relate to:

  - Patient selection (fracture acuity and pain severity)
  - Operator technique (volume of injected cement and/or technique used for injection)
  - Nature of the OPLA
  - Outcome measurement
  - Use of blinding.

  These are discussed in turn below.

  Patient selection

  Fracture acuity

  The Buchbinder and INVEST studies included patients whose fractures were up to 12 months old, as did the studies by Blasco et al and Farrokhi et al. Clark et al. have argued that most VCFs heal within 8 weeks.[107] It would therefore follow that in these four studies, in which the average time since fracture ranged from 9.5 to around 30 weeks, PVP was carried out on fractures which, in most patients, had already healed. This would make it unlikely that vertebral augmentation would have any effect on fracture pain by means of fracture fixation.[112] By contrast, VERTOS II recruited

patients who had pain of no more than 6 weeks' duration,[17] and it has therefore been suggested that it provides the best evidence relating to the effectiveness of PVP in patients with acute osteoporotic VCF.[112] This assertion is misleading. In addition to the lack of blinding in VERTOS II, which reduces the quality of the study, the delay between recruitment and performance of PVP (9.4 days $\pm$8.1) meant that many patients would have pain of more than 6 weeks' duration by the time PVP was performed (see section 5.2.2). Furthermore, aggregation of data from the two blinded placebo-controlled trials showed that outcomes did not differ between those with acute ($\leq$6 weeks) and subacute fractures (>6 weeks).[109] When associations between fracture age and clinical outcomes have been explored in large case series, most of the evidence also suggests no association.[207,208,210] The one exception to this was the study by Ryu et al.[214] who found significant correlations between fracture age and pain, activity, and analgesic use. It was not clear why this discrepancy with the other case series was observed, and the authors seem to have controlled for confounding factors through the use of multiple regression.

Moreover, there is considerable debate about the appropriate timing of PVP and BKP because of evidence that a substantial proportion of VCFs heal without intervention. In VERTOS II, 53% of patients who initially met the inclusion criteria and were willing to participate in the study subsequently became ineligible because their pain score had spontaneously fallen below 5 between screening and randomisation.[17] As noted in section 3.1, a small study by Klazen et al found that, by 6 months, 63% of conservatively-treated patients with acute radiographically-diagnosed VCF reported significant pain relief.[19] They therefore suggest that, to avoid unnecessary interventions, PVP or BKP should only be offered to patients in whom the pain of acute VCF persists for 6 months, but recognise that, during that 6 month wait, a proportion of patients will suffer unnecessary pain and days lost from normal activity.[19] Consequently, studies which include patients with pain of more than 6 weeks duration are likely to be more representative of the group of patients who will be considered for vertebral augmentation in clinical practice than those which are limited to patients with pain of less than 6 weeks duration.

Pain severity

Some authors have suggested that PVP is only effective for patients with more severe pain which is unresponsive to treatment with analgesics.[106,107,329] However, as emphasised throughout this review, pain is open to a number of confounding influences, which makes its reliability as an eligibility parameter questionable. In

addition, the individual patient meta-analysis of the two placebo controlled trials of vertebroplasty failed to demonstrate a between-group difference when pain severity (≥8 or <8) was controlled for as a covariate.[109]

Bone scan methods

Gangi and Clark[112] have argued that the use of plain radiograph for fracture identification in the INVEST study[102] was inadequate. Rather, they suggest that MRI is necessary to identify the presence of marrow oedema and therefore confirm the VCF as the source of pain. Similarly, Whitehouse[94] argues that, unless there is an un-united fracture cleft within the vertebra, confirmed by MRI, the pain which persists past 10 weeks is likely to be multifactorial, as true fracture pain will have been succeeded by mechanical back pain. He also notes that research has suggested that some types of fracture, as seen on initial x-ray, seem to progress, and that cement augmentation is likely to be particularly beneficial in such fractures as it will prevent fracture progression, whereas conservative treatment should be recommended initially in patients whose fracture morphology suggests that it is unlikely to progress, unless they have uncontrollable pain.[94]

Operator technique

The technique used in the Buchbinder study to inject cement (i.e. 13 gauge needles and cement hand-injected using 1-cc syringes[101,150]) has been criticised on the grounds that, to achieve adequate filling of the vertebral body in lumbar fractures, either an 11-gauge needle or a high-pressure injecting system should have been used.[112]

It has also been suggested that the volume of cement injected by Buchbinder et al was too low.[21,112] Aebi has also criticised the INVEST study on this basis, suggesting that, as the mean amount of cement injected in both studies was inadequate, the investigators in essence compared two placebo operations.[21] However, Kaufmann et al[330] noted that greater cement volumes may have better outcomes but higher risks of AEs. Moreover, although Al-Ali et al[331] found a mean volume of cement injected of 5.1±2.2 mL in 600 osteoporotic fractures treated by vertebroplasty, the range was 1.0 to 16.0 mL, and no correlation was observed between the volume of cement and pain improvement. However, they noted that the volume injected was sufficient to fill the intravertebral cleft and, so long as that was done, the volume of cement used was not a determining factor in the degree of pain relief.[331]

Operative placebo with local anaesthesia (OPLA)

Debate is ongoing with respect to the impact of the "sham" procedures in the Buchbinder and INVEST studies. In the unblinded LABEL study, Brinjikji et al.[332] investigated the efficacy for pain relief of injected lidocaine and bupivacaine at the site of painful osteoporotic compression fractures (n=19 consecutive patients presenting for consideration of vertebroplasty between April 2009 and Jan 2010). They compared the changes in the RDQ and average 24-hour pain at days 1 and 3 post-injection with those recorded in blinded control patients from the INVEST lead site (n=16), and found that an unblinded injection of local anaesthetic was ineffective in treating pain from osteoporotic VCFs; significantly greater improvements were seen in the INVEST control patients. This appears to suggest that factors other than local anaesthesia were responsible for that observed improvement. Miller[333] also noted the possibility of high placebo response in any interventional procedure. By contrast, a recent nonrandomised case series[83] found that facet joint injections resolved pain arising in up to a third of patients with VCF who were considered suitable for treatment with PVP, while vertebral augmentation appeared to be clinically effective in patients who failed to respond to the facet joint injections. The authors noted that it was not possible to pre-select those patients who would respond to facet joint injection, but suggested that their results supported the hypothesis that PVP was potentially effective in those patients whose pain arose largely from the VCF itself.

Outcome measurement

A North American Spine Society (NASS) commentary questioned the measurement of pain in both Buchbinder and INVEST, arguing that neither study appeared to make any attempt to assess whether baseline pain was specific to the fracture.[106] These authors suggested that investigators should percuss or palpitate the spinal levels systematically in order to ascertain the area of maximum focal tenderness. However, only three of the open-label trials (Farrokhi,[133] VERTOS,[87] and VERTOS II[18]) reported undertaking such a procedure.[17,147,153] Moreover, as the NASS commentary noted, the improvements observed in the vertebroplasty group of the Buchbinder and INVEST studies were not dissimilar to those observed in unblinded trials. Perhaps a more important issue is the likelihood of confounding in subjective ratings of pain.[83] However, this is an issue in any trial using visual or numeric pain rating scales, and is not specific to Buchbinder and INVEST. As is argued in section 4.1, measures of functional status are more useful than self-reported pain for assessing the impact of vertebral fractures on the patient's daily life because they are more objective; they have therefore been given priority in the results section.

Also related to the issue of pain measurement, Lotz[334] has suggested that, in the INVEST study, the significant difference in crossover rates between treatment and control groups (12% vs. 43%) indicates a degree of patient dissatisfaction with the OPLA procedure which was not fully captured by the pain scales. Indeed, as Doidge et al point out, those who crossed over at one month had worse outcomes for pain and functional status, irrespective of treatment group. This may suggest crossover is a reliable proxy for global effectiveness.[56] However, in an analysis of the crossover data from INVEST, Brinjikji et al[335] noted that baseline pain duration and treatment site were associated with ability to correctly guess treatment allocation in the control group only. That is, poor responders in the control group were able to guess their allocation, whilst good responders in the vertebroplasty group were not. Furthermore, Kallmes et al argue that, as nearly all crossovers occurred after 30 days, they did not affect the primary conclusion that there were no important differences in outcomes between the groups at 1 month.[336]

Length of follow-up presents a further set of interpretive challenges. On the one hand, the benefits associated with vertebral augmentation from the open-label trials were all in the short- to medium-term, with few benefits seen after 6 months. Indeed, this pattern is to be expected in terms of fracture healing and regression to the mean. However, Aebi has suggested that follow-up of 1 year is short to capture the consequences of osteoporotic vertebral compression fractures with increasing kyphosis which may ultimately lead to death.[21] It is possible that the evidence linking vertebral augmentation to improved survival rates[34] may be showing the effect of kyphosis, though this hypothesis is yet to be addressed in clinical trials.

Use of blinding

Perhaps the most convincing reasons put forward for discrepancies between findings from the open-label trials of PVP, and those of Buchbinder and INVEST, relate to the use of blinding. Wood et al,[337] and Psaty and Prentice,[338] have presented empirical evidence that lack of blinding results in an average 25% over-estimate of relative treatment benefit. Exaggerations of effectiveness are also likely to be high in any interventional procedure.[333] Factors such as strong patient and physician expectations of effectiveness, and reconfigurations of meaning within the illness experience, are all likely to play a part in determining the apparent strength of an effect. Indeed, such factors have been shown to have an objective neurophysiological impact on pain pathways.[339,340] This is not to say, therefore, that PVP lacks efficacy per se; but rather

that the mechanisms that lead to improvement may be unrelated to the injection of bone cement.

In addition to the potential influence of the placebo response, the apparent disparity between the OPLA trials and some of the open-label trials may be partly explained by response bias. Miller et al[333] explored this possibility, suggesting that the lack of blinding in the VERTOS II trial may have led to a preponderance among participants in the OPM arm to exaggerate their pain levels due to dissatisfaction with not having received the procedure. Conversely, the participants in the PVP arm may have exaggerated their improvements either due to expectations that they should be getting better after the intervention, or to please the investigators. While it would be difficult to maintain that response bias can account for more objective functional outcomes, it seems probable that the combination of response bias and placebo effects could explain a substantial degree of the inter-study variability.

External validity

Several factors may affect the external validity of the included studies. The first relates to the potential learning curve relating to the vertebral augmentation procedures. It seems reasonable to assume, and indeed is in some cases stated, that the procedures reported in those studies were performed by experienced personnel, and therefore their results may differ from those obtained by less experienced practitioners. There is little evidence to indicate how many procedures a practitioner needs to perform to achieve a high standard. McDonald et al compared the outcomes of PVP performed in the Mayo Clinic, Minnesota, USA, by two interventional neuroradiologists with substantial previous vertebroplasty experience and five experienced interventional neuroradiologists who were initially new to the procedure; the "experienced" operators estimated that they had performed at least 150 PVPs prior to the commencement of the study. Patient outcomes appeared to be broadly similar regardless of operator experience, although loss to follow-up limited exploration of long-term outcomes. However, both the volume of cement used and postoperative pain (as assessed by pain at rest and RDQ scores 1 week after PVP) were higher with 'novice' than with experienced operators, but decreased as the 'novice' operators gained more experience.[341] Thus, although a learning curve can be observed, its effects seem to be relatively limited. However, as the authors caution, the 'novice' operators in their study were all highly skilled interventional radiologists with substantial clinical experience prior to the study, and the results are therefore not necessarily generalisable to less skilled personnel.[341]

There is potentially an issue about the type of cement used in the studies, as newer generation cements (including those which are highly viscous) have been designed to reduce the risk of cement leakage[35]

- Syed conducted a RCT of PVP in patients with osteoporotic VCF comparing PMMA with Cortoss™, a bioactive composite – but this abstract[342] does not report that comparison.
- Blattert et al[343] RCT compared BKP in patients with osteoporotic VCF (including burst fractures) with PMMA and Norian SRS, a calcium phosphate/carbonate cement. Each cement was associated with leaks in 5/30 vertebrae. PMMA was associated with vascular embolism in 2 patients; however, with Norian SRS, there were 9 cases of cement failure, (i.e. at follow-up at 6 weeks, they showed radiographic signs of early cement fracture) all in burst fractures. The investigators therefore did not recommend its use in BKP (there was also persistent haemorrhaging from one vertebral body which partially washed out the cement before it could set – this is less likely to happen with PMMA). (advantage of calcium phosphate-based cements is that they set at a significantly lower temperature than PMMA, and therefore there is less risk of thermal damage to adjacent structures)
- Anselmetti et al[344] in a single-centre RCT compared PVP with standard low-viscosity PMMA and high-viscosity PMMA designed for injection through a proprietary delivery system (Confidence Type I, Disc-O-Tech, Israel) in patients with VCF of any origin; all procedures were performed by the same experienced operator. CT scans were performed 1 hour after PVP to evaluate cement perfusion, leakages, and possible complications; when a venous leak was detected, CT scan of the lungs was performed to assess the possibility of PMMA embolism. No symptomatic cement leaks occurred in either group; asymptomatic venous leaks were significantly less common in vertebrae treated with high-viscosity PMMA than in those treated with low-viscosity PMMA, but the reduction in the number of leaks into the disk was not statistically significant (see Tables 30 and 31).

Table 30: Number of patients with asymptomatic cement leaks[344]

All patients	High-viscosity PMMA	Low-viscosity PMMA	P value
Venous leaks	6/30 (20%)	24/30 (80%)	NR
Leaks into the disk	6/30 (20%)	11/30 (36.6%)	NR
Osteoporotic patients			
Venous leaks	4/23 (17.4%)	19/23 (82.6%)	NR
Leaks into the disk	5/23 (21.7%)	8/23 (34.8%)	NR

Table 31: Number of treated vertebrae with asymptomatic cement leaks[344]

All patients	High-viscosity PMMA	Low-viscosity PMMA	P value
Venous leaks	8/98 (8.2%)	38/92 (41.3%)	<0.0001
Leaks into the disk	6/98 (6.1%)	12/92 (13.0%)	0.1374
Osteoporotic patients			
Venous leaks	6/77 (7.8%)	30/71 (42.3%)	NR
Leaks into the disk	5/77 (6.5%)	9/71 (12.7%)	NR

Summary of key findings

Summary

The included studies measured back-specific functional status using a number of instruments, including the RDQ, SOF-ADL, and indicators of disability such as walking aids. FREE reported significantly better RDQ outcomes in the BKP group at 1 and 12 months, although the 95% CIs included the possibility of a lack of MCID at both time points. VERTOS II reported a significant difference favouring PVP at 12 months. However, the two blinded placebo-controlled studies found no statistically significant between group differences. Farrokhi et al measured functional status with an instrument based on the ODI, and found a statistically significant difference favouring PVP at all follow-up time points.

Five studies (Buchbinder,[101] FREE,[151] INVEST,[102] Rousing,[185] and VERTOS II[17]) measured quality of life using the EQ-5D. However, VERTOS II did not report follow-up values, and both Rousing and Buchbinder only began to collect EQ-5D data part way through the trials. Aggregation of IPD from the Buchbinder and INVEST trials[109] found no significant difference in EQ-5D in short- or medium-term outcomes. Conversely, the FREE trial found significant differences favouring BKP throughout follow-up. Another commonly used quality of life measure was QUALEFFO, which was reported in four studies (Blasco,[146]

Buchbinder,[101] VERTOS,[153] and VERTOS II[17]). Broadly speaking, there was a tendency toward favouring PVP. However, the OPLA trial found no significant between group differences on any of the QUALEFFO subscales at any time point.

Four open-label studies (Farrokhi,[147] FREE,[151] Rousing,[185] and VERTOS II[17]) found statistically significant differences between groups in short- and medium-term improvement in pain. Conversely, the two double-blinded, placebo-controlled studies of vertebroplasty, Buchbinder[101] and INVEST,[102] found no statistically significant between-group differences in pain, and the trend towards a higher rate of clinically meaningful improvement in the INVEST trial may be confounded by greater opioid use in the PVP group. INVEST also measured pain frequency and pain bothersomeness on a scale of 0 to 4, and did not find a statistically significant, or clinically meaningful, difference between groups. A similar picture emerged for analgesic use. INVEST and Buchbinder reported reductions in opioid use from baseline in both the PVP and control group, with no statistically significant between-group differences at any time point. However, VERTOS and VERTOS II reported greater short-term reductions in analgesic use among patients treated with PVP, while short- and medium-term differences favouring BKP were reported in the FREE trial.

None of the studies reported statistically significant differences in mortality between treatment groups. However, none of the studies were powered to detect this outcome. A meta-analysis was performed on the three studies which reported all-cause mortality at 12 months.[17,146,185] Although the pooled result slightly favoured PVP, the effect failed to reach statistical significance. Furthermore, there are plausible biomechanical explanations as to why vertebral augmentation may increase life expectancy; these include improvement of lung function due to correction of kyphotic deformity,[345] and mitigation of impaired physical function through pain relief.[345,346] However, the absence of randomisation in this cohort means that confounding factors cannot be ruled out. A formal critique of mortality data from observational databases (Appendix 12) concludes that ‘it is possible that there is a causal difference in mortality between patients treated using OPM and OP patients given the size of the effect. Appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. It is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment based on the data presented in the studies included here.’ There is also considerable uncertainty were BKP and PVP assumed to have a mortality benefit, in whether OPLA would also produce a mortality benefit. However, there were no data on this.

PVP and BKP appear to be reasonably safe procedures, with a low rate of intra-and postoperative complications. However, when complications do arise from vertebral augmentation, they can be serious. Among the several large case series analysed in this review, one reported a death, though it seemed likely that factors other than the vertebral augmentation played a part.[207] However, several deaths directly related to augmentation procedures have been noted in case reports. Cement extrusion was common, including extrusion into the epidural space. However, most reported extrusions remained asymptomatic. Vertebral augmentation does seem to be associated with a higher risk of new adjacent fractures. A substantial number of case reports of pulmonary embolism, and several case reports of rare but potentially serious complications were identified.

Discussion of potential subgroups

The included studies were too small to permit the identification of subgroups of patients who might benefit from PVP or BKP. While there has been considerable debate on whether vertebral augmentation is more effective in the treatment of acute or chronic fractures,[334,336,347,348] analysis of individual patient data from the two double-blinded RCTs of PVP suggested that effectiveness was unrelated to fracture acuity.[109]

It has been suggested that PVP and BKP are more successful in patients with a mobile pseudarthrotic cleft pattern of fracture than in those with the more common non-mobile fracture,[202] but further research is required to explore this possibility. It has also been suggested that there is a substantial subgroup of patients whose pain is not a direct result of VCF, but of overload of facet joints, paraspinal muscles and impingement of spinous processes; such patients respond to facet joint injection while those who do not respond to this treatment have an excellent response to vertebroplasty.[83] However, further evidence from clinical trials would be required to confirm the importance of such subgroups.

Conclusions for clinical effectiveness

PVP and BKP perform significantly better than OPM in reducing pain and disability, and improving HRQoL. There is some evidence that PVP and BKP may lead to reductions in mortality; however, this effect has not yet been confirmed in clinical trials with a randomisation procedure, so the causal mechanisms remain unclear. As yet, there is no convincing evidence that PVP and BKP perform better than blinded administration of local anaesthetic to the affected area in terms of reducing pain and improving HRQoL. However, this may be due, at least in part, to inadequate patient selection methods with some patients (who cannot be identified a priori) receiving considerable benefit. Although the incidence of severe complications arising from vertebral augmentation is low, leakage of cement into the epidural space can pose a serious risk to health, and a small number of procedure-related deaths have been noted in previous case reports.

6. ASSESSMENT OF COST-EFFECTIVENESS

6.1 Systematic review of existing cost-effectiveness evidence

6.1.1. Previously published economic models of PVP and BKP

From the literature review 243 potential data sources were identified, with full copies of 43 requested. Only one mathematical model assessing the cost-effectiveness of BKP or PVP in the defined population was found. This concurred with the conclusions presented by Medtronic. The identified manuscript was authored by Strom et al.[349] This used a markov cohort methodology to ascertain the cost-effectiveness of BKP compared with OPM. The model simulated the experiences of hypothetical patients until death or age 100, with EQ-5D scores taken directly from the FREE study.[151] It was assumed that the EQ-5D scores would be independent of intervention 3 years post BKP or OPM with a linear decline between 12 months and 36 months. The risks of future vertebral fracture and the risks of mortality after vertebral fracture were incorporated.

The base case assumed a cohort of 70 year old women and men with a T-Score of -2.5SD and estimated that BKP would be associated with an additional cost of £1494 to obtain 0.169 QALYs at a ratio of £8840 per QALY gained.

The model of Strom et al was updated to include PVP as an intervention and incorporate the potential beneficial effect of BKP and PVP on mortality and used within the Medtronic submission. As such, it is deemed that the results presented have been superceded.

6.1.2. The models submitted by the manufacturers

The Johnson and Johnson model

Johnson and Johnson submitted a de novo cost-effectiveness model to determine the cost effectiveness of PVP, BKP, OPLA (denoted as 'invasive control procedure' and also as 'sham') and OPM (termed 'non-invasive management'). The perspective of the analysis was that of direct NHS and personal and social services costs. In the base case the time horizon was that of one year, with discounting of both costs and benefits at 3.5% per annum in sensitivity analyses extending beyond a one-year time horizon.

The base case assumed a one year time horizon assuming that all benefit was lost at one year, excluded OPLA as a comparator and included clinical evidence from all relevant studies identified (Buchbinder;[101] Chen;[350] Chen;[351] Farrokhi;[147] Klazen;[17] Kallmes[102]; Liu;[352] Rousing;[185] and Wardlaw[151]) A 'target population' was also denoted where only the results from Chen;[351] Klazen;[17] Liu;[352] Rousing[185] and Wardlaw[151] were included as these were reported as only including fractures within the previous three months. The Assessment Group note however, that no differential effects of vertebroplasty compared with placebo was

observed when the duration of pain was divided into≤ 6 weeks and ˃ 6 weeks categories.[109] Eight alternative scenarios were evaluated, which are detailed fully in the Johnson and Johnson submission (pages 136-137 and cross-references).

1) Incorporating data from the OPLA trials
2) Incorporating data from the OPLA trials but assuming that these could be pooled with OPM
3) Extending the time horizon to beyond one year
4) As 3) but using target population results
5) Using an alternative bottom up costing methodology and payment by results tariff
6) As 5) but using target population results
7) Using direct EQ-5D values directly
8) As 7) but using target population results

Within the model patients are assigned a VAS score at baseline and then at 2 weeks dependent on the intervention received. The treatment dependent VAS is updated at 1 month, 6 months and 12 months. The values for treatment dependent VAS were estimated from a network meta-analysis. For information, Table 78 of the Johnson and Johnson submission, for Scenario 1, which includes OPLA trials, is replicated in Table 32; other analyses are contained within their submission.

Table 32: The VAS scores from Johnson and Johnson's network meta analysis including OPLA

Treatment	VAS over time, mean (95% CI)			
	2 weeks	1 month	6 months	12 months
Vertebroplasty	3.360 (2.810–3.900)	2.530 (1.430–3.630)	2.410 (1.880–2.940)	2.170 (1.570–2.780)
Kyphoplasty	3.650 (3.10 - 4.190)	2.990 (1.780–4.200)	2.420 (1.880–2.960)	2.830 (2.220–3.440)
NIM	5.920 (5.530–6.310)	4.940 (3.980–5.900)	4.110 (3.820–4.410)	3.810 (3.530–4.080)
OPLA (Invasive control procedure ("sham"))	3.090 (2.180–4.020)	3.080 (1.720–4.440)	2.410 (0.860–3.970)	Not available – assumed equal to the 6 month value

Abbreviations: CI, confidence interval; NIM, non-invasive management; VAS, visual analogue scale.

The Assessment Group makes two comments regarding the network meta-analysis conducted by the manufacturer. Firstly, there was no attempt to extrapolate or interpolate data from RCTs if they did not report VAS scores at the designated time intervals, this could cause discrepancy within the longitudinal data. Secondly, a further trial, Blasco et al[146] was published after the completion of the manufacturer's systematic review. This trial had similar VAS scores for both PVP and OPM, with both values being relatively high. If the manufacturer had included this study, the VAS scores in all arms would have increased and the relative difference between OPM and both PVP and BKP be reduced.

A regression analysis was conducted to translate VAS scores into utility from which quality adjusted life years (QALYs) can be calculated. The formula was EQ-5D = 0.9242 – 0.0955 * VAS score. Co-variance between the intercept and the slope did not appear to be incorporated.

The submission undertook analyses on potential adverse events looking specifically at cement leakage and refracture rates. For full details refer to the manufacturer's submission

It was concluded that cement leakage was highest using low-viscosity cements, and that cement leakage using high-viscosity cements was equivalent to that in BKP. Table 71 of the manufacturer's submission is reproduced in Table 33. It is commented that the pooled odds figure relates to the odds that a patient will experience an event if receiving a treatment rather than being odds ratios.

Table 33: Odds for cement leakage by treatment

Treatment	Pooled odds	SE
NIM	0	0
Vertebroplasty LV	0.814	0.776
Vertebroplasty HV	0.167	1.385
Unilateral/unipedicular kyphoplasty	0.131	1.149
Bilateral/bipedicular kyphoplasty	0.074	1.679

Abbreviations: HV, high viscosity; LV, low viscosity; NIM, non-invasive management; SE, standard error.

Regarding fracture rates it was stated that it could not be concluded that there was a significant difference between any of the treatments in terms of re-fracture rates. Table 72 of the manufacturer's submission is reproduced in Table 34. It is commented by the Assessment Group that these values include all fracture rates and that the conclusions may differ if only adjacent fractures were considered.

Table 34 Comparison of refracture rates by intervention

Treatment comparison	Mean HR	Median HR	Lower 95% Credible Interval	Upper 95% Credible Interval
Vertebroplasty LV vs. NIM	0.63	0.61	0.34	1.07
Kyphoplasty vs. NIM	1.24	1.20	0.71	2.03
OPLA (Invasive control procedure ("sham")) vs. NIM	1.21	0.80	0.15	4.70
Vertebroplasty HV vs. NIM	13.41	1.49	0.11	48.8
Vertebroplasty Cortoss vs. NIM	0.56	0.52	0.23	1.15
Vertebroplasty CHC vs. NIM	7.57	0.76	0.05	26.6

The assumed acquisition costs

The list prices for PVP using the CONFIDENCE SPINAL CEMENT SYSTEM™ were taken from the Johnson and Johnson submission[35]. The costs vary according to the number of levels that need to be treated, and reported to be £1358 at one level, £1784 at two levels and £1848 for a three level approach. It is noted that Table 83 of the Johnson and Johnson submission appears to contradict the text regarding the cement required in the two level procedure, with the text stating that 11cc would be required but calculate the costs assuming 7cc were sufficient. The distribution of operations between one-, two- and three- levels were extracted for Johnson and Johnson by Dr Foster, and are 58.9%, 20.5% and 20.5% respectively. The costs per level were multiplied by these proportions to arrive at a weighted cost of £1472. If 11cc of cement were assumed in the two-level operation rather than 7cc this value would increase to £1546.

The Johnson and Johnson submission inflated the price of a BKP operation reported in Strom et al[349] and assumed a cost of £2842 for BKP. The manufacturer also assumed the cost of OPLA would equal the cost of PVP; this may be questionable particularly when it is assumed that high viscosity cement would be used in the OPLA rather than cheaper low viscosity cement.

The costs of the preliminary phase, the operating phase and the post-operative phase have previously been reported in Ström O et al.[349] These were inflated to 2009/10 prices within the Johnson and Johnson submission, and are partially replicated in Table 35. The table in the Johnson and Johnson appears to misreport the operating room costs, which should be £275, as used in their mathematical model. This has been amended in Table 35. It was assumed that these costs are applicable to both BKP and PVP.

The Assessment Group comment that there appeared to be a typographical error in the manufacturer's mathmatical model in which only 10% of patients receiving BKP were assumed to consume operating room resources; it was assumed that this value was intended to be 100%. As such the overall cost-effectiveness results are likely to be favourable to BKP.

Johnson and Johnson also undertook bottom-up costing within the operating phase. This approach used data on procedure times and number of vertebral levels treated, collected using a bespoke iPad app designed to measure the total duration of the operating room episode in minutes. These data were collected at five hospitals (details confidential). Data from the iPad app were used in conjunction with data from an audit of vertebroplasty procedures obtained from two hospitals (identity confidential) currently offering vertebroplasty to generate estimated costs for the procedure. Whilst the average weighted operating cost per procedure can be made public the breakdown of constituent parts remain confidential. Data obtained included the cost and volumes of surgical consumables, medication costs (including sedation and antibiotics), theatre costs and staff costs." The data from this analysis are replicated in Tables 36 and 37.

Table 35 Resource use in the preliminary, operating and post-operative phases.

	Cost per hour	No. of hours	Cost of resource	% of pts	Cost
Preliminary phase					
Surgeon	£106	0.25	-	-	£26
Radiologist	£85	0.50	-	-	£42
Nurse	£16	1.00	-	-	£16
Spine X-ray	-	-	£76	100%	£76
MRI	-	-	£275	100%	£275
ECG	-		£68	100%	£68
Blood test(s):	-	-	£21	100%	£21
Pain therapy	-	-	£16	100%	£16
Sum of preliminary phase					**£540**
Operating phase					
Anaesthetist	£106	1.00	-	-	£106
Nurse, anaesthesiology	£12	1.00	-	-	£12
Drugs	-	-	£38	100%	£38
Surgeon	£106	1.00	-	-	£106
Radiographer	£32	1.00	-	-	£32
Nurse, surgery	£17	2.00	-	-	£34
Consumables	-	-	£95	100%	£95
Operating room	-	-	£275	100%	£275
Sum of operating phase					**£698**
Post-operative phase					
Nurse	£4.8	24.00	-	-	£114
Surgeon	£106	0.50	-	-	£53
Spine X-ray	-	-	£76	100%	£76
Sum of post-operative phase					**£243**
Total Sum					**£1479**

Abbreviations: ECG, electrocardiogram; MRI, magnetic resonance imaging; pts, patients. The costs presented in this table are derived from Ström 2010 and inflated to 2009/10 values. Costs may not tally exactly due to rounding of numbers

Table 36: Average procedure length and split of levels

Number of levels	**% split**	**Average procedure duration (mins)**
1 level	59%	31.75
2 or more levels	41%	46.20
Weighted average duration (mins)		**37.69**

Table 37: Estimated costs associated with vertebroplasty

Resource based costs (per hour)	**Hourly rate (inc. on costs)**	**% of pts**	**Av. cost per procedure**
Consultant	[redacted]	[redacted]	[redacted]
Anaesthetist	[redacted]	[redacted]	[redacted]
Theatre staff (Includes 3 theatre staff)	[redacted]	[redacted]	[redacted]
Radiographer	[redacted]	[redacted]	[redacted]
Recovery	[redacted]	[redacted]	[redacted]
Theatre session (per hour)	[redacted]	[redacted]	[redacted]
Sedation	[redacted]	[redacted]	[redacted]
General anaesthetic	[redacted]	[redacted]	[redacted]
Surgical consumables	[redacted]	[redacted]	[redacted]
Average weighted cost per procedure			**£527.55**

The costs associated with hospitalisation stay

The length of stay following each intervention.

In the Johnson and Johnson submission, a third party (Dr Foster Intelligence) was employed to extract data based on the ICD code (M80*) and the OPCS 4.5 code (V444 for PVP, V445 for BKP and blank for those treated with OPM).

The assumed hospitalisation costs per day.

Johnson and Johnson assumed a cost of £232 per day based on the Payment by Results national tariff price for an excess bed day associated with VP/KP and NIM healthcare resource group (HRG) codes (HRGs HC04C, HC05C and HD36C).[353] These values are summarised in Table 38.

Table 38 The assumed length of stay, cost per day and total cost assumed in the Johnson and Johnson submission

	Johnson and Johnson		
Intervention	Length of Stay in days (standard error)	Cost per day	Total Cost
PVP	3.24 (0.49)	£232	£752
BKP	4.48 (0.89)	£232	£1039
OPM	12.61 (0.27)	£232	£2926

Johnson and Johnson assumed that OPLA incurred the same hospitalisation costs as PVP

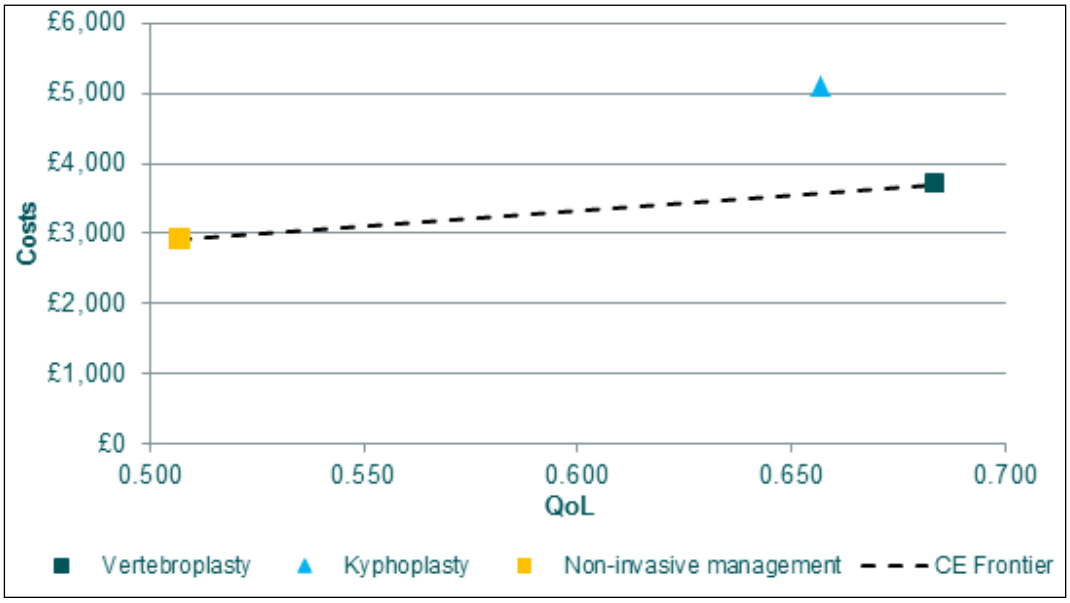
Results presented by Johnson and Johnson

The base case deterministic results are replicated in Table 39. It is seen that PVP was shown to dominate (that is producing more QALYs at a lower cost) BKP. The cost-effectiveness plane and cost-effectiveness acceptability frontier are shown in Figure 16. The incremental cost effectiveness ratio (ICER) between PVP and OPM was £4392 per QALY gained. Explicit comparison between the PVP and BKP results indicated that the ICER between PVP and BKP was 99.86% likely to be below £20,000 per QALY gained.

Table 39 The base case deterministic results in the Johnson and Johnson submission

Treatment	Costs	QALYs
PVP	£3,702	0.684
BKP	£5,113	0.656
OPM	£2,926	0.507

Figure 16 The cost-effectiveness plane and the cost-effectiveness frontier associated with the base case deterministic results in the Johnson and Johnson submission



The cost-effectiveness frontiers and ICERs for the Target Population and the alternative scenarios undertaken by Johnson and Johnson are summarised in Table 40.

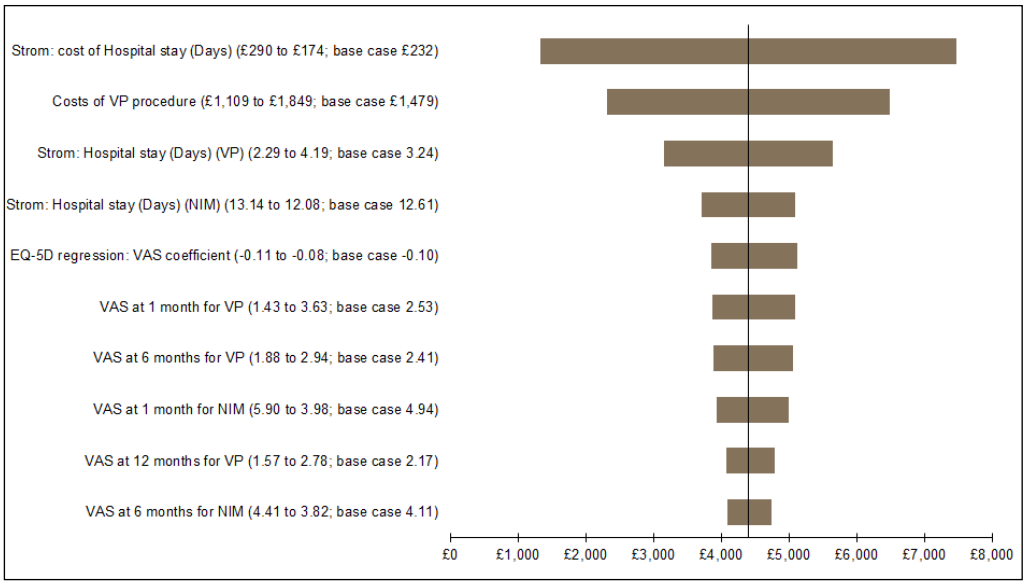
Table 40 The base case deterministic results in the scenario analyses within the Johnson and Johnson submission

Scenario		Cost-effectiveness Frontier	ICER on cost-effectiveness frontier
-	Target Population	OPM / PVP	£4755
1	Incorporating data from the OPLA trials	OPM / PVP	£4392
2	Incorporating data from the OPLA trials but assuming that these could be pooled with OPM	OPM / PVP	£4982
3	Extending the time horizon to ten years with decline in benefit across time	OPM / PVP	£1054
4	As 3 but using target population results	OPM / PVP	£1168
5a)	Using an alternative bottom up costing methodology and payment by results tariff	PVP	-
5b)	Using payment by results tariff	OPM / PVP	£13,595
6a)	As 5a) but using target population results	PVP	-
6b)	As 5b) but using target population results	OPM / PVP	£14,718
7	Using direct EQ-5D values directly	OPM / PVP	£5516
8	As 7 but using target population results	OPM / PVP	£5516

Univariate sensitivity analyses

Johnson and Johnson undertook Univariate sensitivity analysis comparing PVP with OPM and PVP with BKP. The analyses for PVP and OPM are reproduced in Figure 17. The analyses comparing PVP and BKP are not reproduced as in all analyses undertaken PVP dominated BKP

Figure 17 A tornado plot of univariate sensitivity comparing PVP with NIM in the base case



A threshold analysis was undertaken on the parameters contained in Figure 17 in the base case. These values are reproduced in Table 41. Similar analyses were presented for the target population with broadly similar results.

Table 41 A threshold analysis on key parameters affecting the deterministic results in the Johnson and Johnson submission

Variable	Base case	£20,000 / QALY	£30,000 / QALY
Ström: cost of Hospital stay (Days)	£232	-£63†	-£251†
Costs of VP procedure	£1,479	£4,239	£6,008
Ström: Hospital stay (Days) (VP)	3.24	15.14	22.76
Ström: Hospital stay (Days) (OPM)	12.61	0.71	-6.91
EQ-5D regression: VAS coefficient	-0.10	-0.02	-0.01
VAS at 1 month for VP	2.53	8.79	9.38
VAS at 6 months for VP	2.41	5.56	5.86
VAS at 1 month for OPM	4.94	-1.32†	-1.91†
VAS at 12 months for VP	2.17	7.95	8.49
VAS at 6 months for OPM	4.11	0.96	0.66

Abbreviations: EQ-5D, European Quality of Life-5 Dimensions; KP, kyphoplasty; NIM, non-invasive management; QALY, quality-adjusted life year; VAS, visual analogue scale; VP, vertebroplasty. †indicates an illogical parameter value.

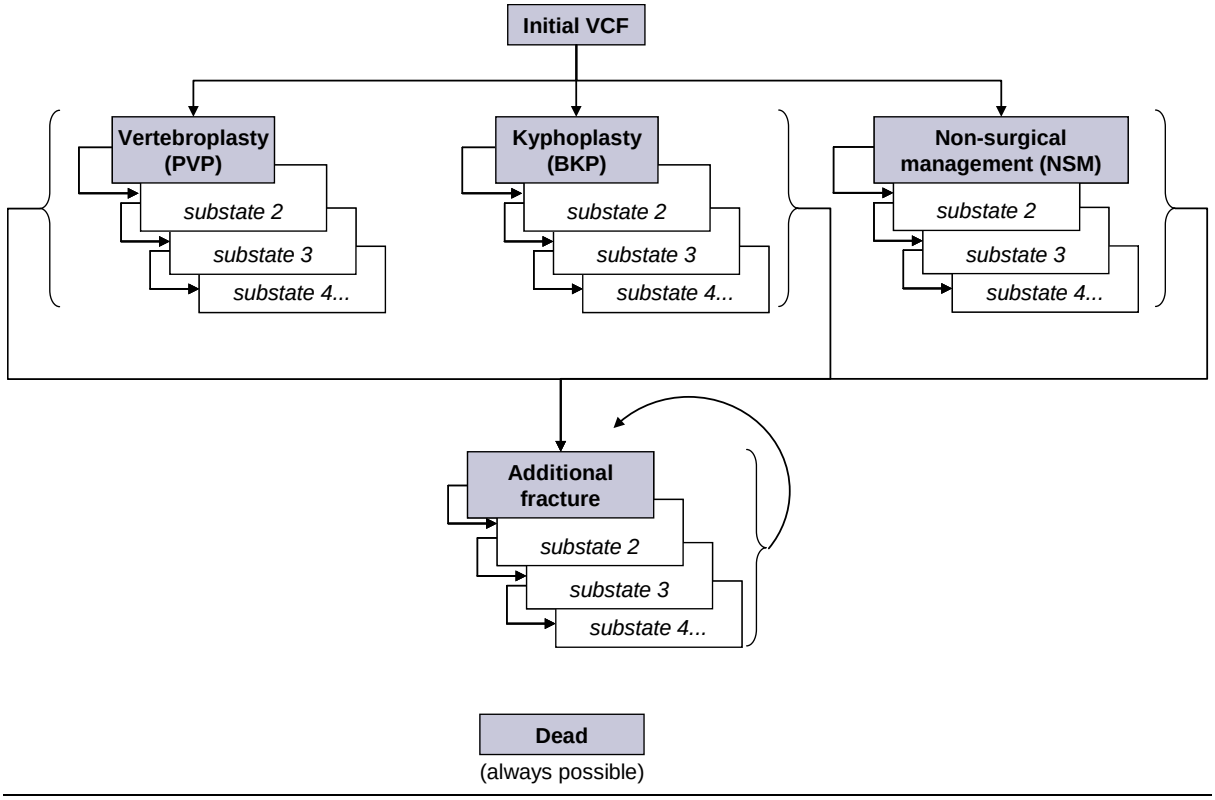
Probabilistic Results

The results from the probabilistic analyses were very similar to those for the deterministic base case. PVP was still estimated to dominate BKP, with the ICER between PVP and OPM being £4388 per QALY gained, compared with £4392 deterministically. The comparison between the PVP and BKP results indicated that the ICER between PVP and BKP was 98.55% likely to be below £20,000 per QALY gained. As the individual cost and QALY components were very similar in both the probabilistic and deterministic analyses it was deemed that the model was linear, and for brevity, only the deterministic values have been reported.

The Medtronic model

Medtronic submitted a Markov Tunnel model using a patient lifetime approach. The tunnel approach allows the time in a health state to be reflected in model parameters such as transition probabilities, costs and utilities. A time cycle of six months was used and a NHS perspective employed. Both health and costs were discounted at 3.5% per annum. The diagram of the model presented by Medtronic is replicated in Figure 18. The objective of the model was to determine the cost-effectiveness of PVP, BKP, and OPM (termed non-surgical management)

Figure 18 The diagrammatic representation of the Medtronic model.



Medtronic assumed that the patient population was those hospitalised for vertebral fracture, the stated rationale being that the FREE Trial[151] which is the pivotal BKP trial was conducted in hospitalised patients and that BKP is predominantly an inpatient procedure[151] in the UK. In the base case it was assumed that patients were 70 year old with a T-Score of -3.0SD, which was reported to be commensurate with the data within the FREE Trial and Vertos II[17]
The patient remained in their initial treatment health state (progressing through the sub-states) until an additional vertebral fracture occurred or a patient dies. For all patients a subsequent vertebral fracture was assumed to be treated using non-surgical management. The transition probabilities for further vertebral fractures were calculated from equations that are a function of the patient's BMD compared with that of a young woman, age, previous fracture status and the imputed ratio between hip and vertebral fractures at each age, assuming that the Swedish ratio was applicable to the UK. The transition probabilities to death used data from the Human Mortality Database for UK patients[354] and the relative risks of mortality reported in Strom et al[349] for people with a prior vertebral fracture.

The health utility for BKP and OPM were taken directly from the FREE trial.[151] The utility for PVP was estimated assuming that the difference between PVP and OPM reported in the VERTOS II trial [17] could be directly added to the OPM scores in the FREE trial. As the QALY data for VERTOS II were presented only at baseline, 1 month and 12 months the manufacturer inferred the average utility across the one year time horizon. The estimate of the undiscounted QALYs gained in the first year for each treatment is provided in Table 42
It was assumed that the difference in utility between BKP and OPM would linearly decline across one year such that there was no difference three years after the intervention. For PVP it was assumed that the utility after the first year (which was not recorded) would progress similarly for that for BKP. It was assumed that the utility of patients would decline after two years in accordance with population norm data. The source for these data, as apparent from the mathematical model was Ara et al.[355]

Table 42 The assumed undiscounted QALYs gained per treatment in the Medtronic submission base case

	Intervention		
Intervention	OPM	BKP	PVP
0-6 months	0.219	0.276	0.273
6-12 months	0.255	0.311	0.309
13-18 months	0.260	0.307	0.305
18-24 months	0.265	0.307	0.305
24-30 months	0.264	0.292	0.291
30-36 months	0.264	0.278	0.277
36-42 months	0.263	0.263	0.263

The model assumes that both BKP and PVP are associated with a mortality benefit compared with OPM. The relative risk for BKP was set at [redacted] and for PVP was set at [redacted]. The manufacturer notes that these values have since been updated, but these data became available too close to their submission date to incorporate within the model.

No adverse events were included in the model bar recurrent fracture, with lack of data being the reported reason for the omission, although the submission does state that associated consequences may be 'substantial'. The rate and consequences of additional fractures were assumed independent of treatment as neither FREE nor VERTOS II detected a significant difference in the incidence of new fractures amongst treatments.

The assumed acquisition costs

The list price for a BKP kit (£2600.50) has been taken from the Medtronic submission,[34] Medtronic additionally quote a lower price as an average selling price but this value (£1900). An additional £96 has been added for devices used in the operation.

Medtronic assume a cost of PVP of [redacted] reported to be the average selling price of De Puy's Spine PVP plus an additional £53 for devices used in the operation.

The costs associated with the operation

The costs of the preliminary phase, the operating phase and the post-operative phase have previously been reported in Ström O et al.[349] Medtronic updated these costs in Table 49 of their submission, which are replicated in Table 43.

Table 43 Procedure Costs reported in the Medtronic Submission

	BKP	PVP	NSM
Procedure costs			
Devices	96	53	0
Consumables	1,900	[redacted]	0
Other procedure costs			
Preliminary Phase			
Interventional Radiologist	0	107	0
Surgeon	107	0	0
Nurse	16	18	0
Rx Spine	77	77	0
MRI	176	176	0
ECG	68	68	0
Blood Test	21	21	0
Drugs	16	16	0
Operating Phase			
Anesthetist	107	107	0
Nurse - Anesthesia	12	13	0
Drugs	38	22	0
Radiologist	0	107	0
Surgeon	107	0	0
Nurse - Operation	17	17	0
Cost of operating room	160	160	0
Post operative phase			
Nurse	41	41	0
Drugs	27	63	0
Total Procedure costs	**2,986**	[redacted]	**0**

The costs associated with hospitalisation stay

The length of stay following each intervention.

In the Medtronic submission the length of stay were reported to be taken from Hospital Episode Statistics 2010/11 data. These values are summarised in Table 44.

The assumed hospitalisation costs per day.

In the Medtronic submission the assumed cost per day in hospital was taken from NHS Ref costs 2009/10/11, and was £457.

Table 44 The assumed length of stay, cost per day and total cost assumed in the Medtronic submission

Intervention	Length of Stay in days	Cost per day	Total Cost
PVP	6.2	£457	£2833
BKP	5.1	£457	£2331
OPM	9.5	£457	£4342

The results presented by Medtronic

Deterministic Results

Medtronic presented both deterministic and probabilistic results for OPM, PVP and BKP; OPLA was not considered a comparator. Medtronic estimated that all three treatments lay on the cost-effectiveness frontier. In the deterministic analysis the ICER between OPM and PVP was £2053 per QALY gained, whilst that for BKP compared with PVP was £2510.

Table 45 The deterministic base case results presented in the Medtronic submission

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus NSM (QALYs)	ICER (£) incremental (QALYs)
Deterministic analysis								
OPM	5,394	9.851	4.976					
PVP	6,112	10.113	5.325	718	0.26	0.35	2,053	2,053
BKP	6,403	10.319	5.441	1,008	0.47	0.47	2,167	2,510
Probabilistic analysis								
OPM	5,394		4.975					
PVP	6,132		5.327	738	0.00	0.35	2,100	2,100
BKP	6,385		5.443	991	0.00	0.47	2,118	2,174

Sensitivity analyses were conducted by Medtronic on: the time horizon (Table 25 of Medtronic's submission); the discount rate for costs (Table 26); the discount rates for QALYs (Table 27); the proportion of health utility benefit from the pivotal trial (Table 28); the health utility offset time (Table 29); post fracture mortality rates (Table 31); the price of PVP compared with BKP (Table 33); the unit costs per bed day (Table 34); the assumed T-Score of the cohort (Tables 37 and 38); the age of the cohort (Tables 39 and 40); the removal of bisphosphonate treatment (Table 41); and assuming that all patients were male (Table 42). In each instance the conclusion that BKP produced most QALYs had an ICER below £15,000 per QALY gained compared with either PVP or OPM remained constant. The Assessment Group comment that the results in Table 28 may lack face validity as the OPM QALY value increased when the benefits of the trial rose from 25% to 50% but remained constant when the benefits were assumed to increase from 50% to 75%.

The fact that BKP had an ICER of £12,353 per QALY compared with PVP when the price of PVP was set to zero (Medtronic Table 33) highlights that the assumed mortality effect (which was more favourable to BKP than PVP) was a key driver of the cost-effectiveness results as seen in Table 30 of Medtronic's submission (reproduced below in Table 46). When it was assumed that there was no mortality benefit associated with either PVP or BKP then the ICER of BKP compared with PVP was £27,340 per QALY gained. It is noted that the ICERs for both PVP and BKP compared with NSM remained low, with the key change being the ICER between BKP and PVP.

Table 46 Sensitivity Analyses presented in the Medtronic submission on the impact of assumed mortality benefit.

Technology	Total Costs	Total QALYS	ICER - PVP vs. NSM	ICER - BKP vs. NSM	ICER - BKP vs. PVP
0% mortality benefit					
Deterministic Analysis					
NSM	5,394	4.976	3,245	4,325	27,340
PVP	6,094	5.191			
BKP	6,371	5.201			
50% mortality benefit					
Deterministic Analysis					
NSM	5,394	4.976	2,511	2,881	4,562
PVP	6,103	5.258			
BKP	6,387	5.320			
75% mortality benefit					
Deterministic Analysis					
NSM	5,394	4.976	2,258	2,472	3,233
PVP	6,107	5.292			
BKP	6,395	5.380			

The sensitivity analysis conducted by Medtronic on the assumed length of hospital stay following BKP (Table 32 in the Medtronic submission) also increased the ICER of BKP compared to PVP to over £20,000 per QALY gained. The length of stays for OPM and PVP were maintained at 9.5 days and 6.2 days respectively whilst BKP was increased from the base case of 5.1 days.

Table 47 Sensitivity Analyses presented in the Medtronic submission on the impact of assumed length of stay following BKP

Technology	Total Costs	Total QALYS	ICER - PVP vs. NSM	ICER - BKP vs. NSM	ICER - BKP vs. PVP
7.65 days					
Deterministic Analysis					
NSM	5,394	4.976	2,053	4,670	12,572
PVP	6,112	5.325			
BKP	7,568	5.441			
10.2 days					
Deterministic Analysis					
NSM	5,394	4.976	2,053	7,174	22,634
PVP	6,112	5.325			
BKP	8,733	5.441			

A further sensitivity analysis was conducted assuming no benefit beyond those seen in the FREE[151] and VERTOS II[17] trials (Table 36 of the Medtronic submission). This is partially replicated in Table 48, with the Assessment Group amending the table to correctly implement extended dominance. It is commented that the trials were not directly comparable as FREE reported EQ-5D values for two years, whereas VERTOS II reported values only for one year.

Table 48 Sensitivity Analyses presented in the Medtronic submission on the impact of assuming no further benefit beyond the time horizon of the trials.

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus NSM (QALYs)	ICER (£) incremental (QALYs)
Deterministic analysis								
NSM	5,394	9.85	4.98					
PVP	7,602	9.85	5.08	2,208	0.00	0.11	20,881	Extendedly Dominated
BKP	8,381	9.85	5.17	2,987	0.00	0.19	15,655	9,160

Probabilistic Results

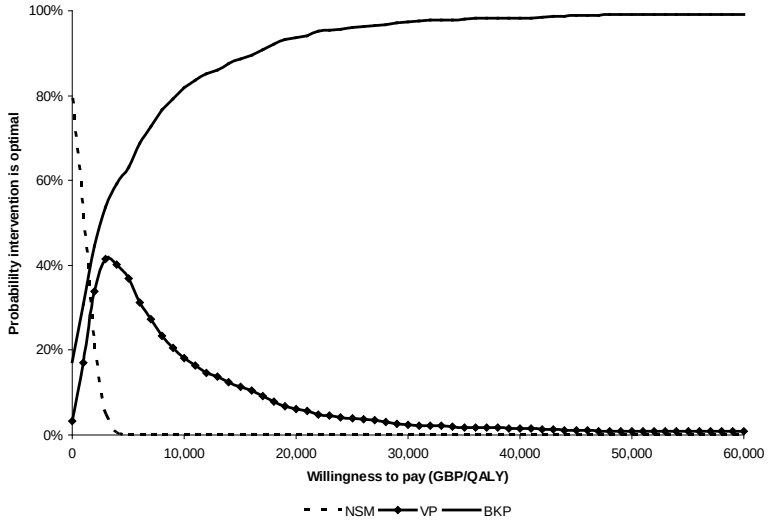
The probabilistic results are replicated in Table 49. When compared with the results in Table 45 is seen that deterministic results and probabilistic results were similar. The Assessment group note that it is unclear why results for total life years gained (LYG) were not reported in the probabilistic analyses.

Table 49 The probabilistic base case results presented in the Medtronic submission

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus NSM (QALYs)	ICER (£) incremental (QALYs)
Probabilistic analysis								
OPM	5,394		4.975					
PVP	6,132		5.327	738	0.00	0.35	2,100	2,100
BKP	6,385		5.443	991	0.00	0.47	2,118	2,174

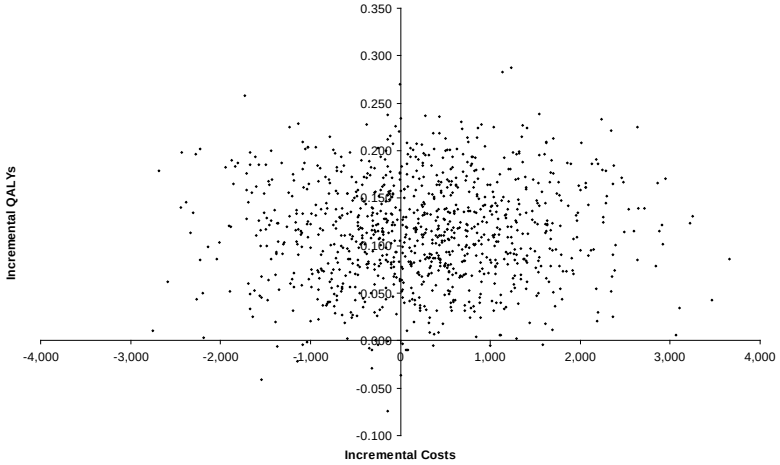
A cost-effectiveness acceptability curve (CEAC) was submitted by Medtronic, which is reproduced in Figure 19.

Figure 19 The CEAC presented in the Medtronic submission



A further comparison of the results between BKP and PVP was provided (and replicated in Figure 20). This indicated that in the large majority of cases BKP provided more QALYs than PVP, with an approximately even distribution between whether BKP or PVP was the more expensive procedure.

Figure 20 A scatter plot of the paired BKP and PVP results.



6.2. The Assessment Group model

The decision problem

Whilst in principle the decision problem appears straightforward comparing PVP (with different cement viscosities), BKP and OPM, in reality, the addition of OPLA as a potential comparator adds significant controversy to the evaluation.

As detailed in the clinical review chapter, and analysed further through the means of a network meta-analysis within this chapter, the results in terms of difference in the level of patient benefit between PVP and OPLA is considerably less than the difference in change in the level of patient benefit between PVP and OPM. This indicates that at least at part of the response for PVP (and through the network of evidence, also therefore BKP) compared with OPM in the open label trials appears to be placebo driven.

The decision to include or exclude OPLA as a comparator can be criticised regardless of the actual conclusion. If OPLA is included then there will be criticism that the use of OPLA treatment within the NHS could, in itself, be deemed unethical and a non cost-effective use of scarce resources. There may be an additional issue regarding whether the components of OPLA, which include providing local anaesthesia and a small incision in the back, meet the criteria for non-invasive management which is the comparator in the scope.[356]

If OPLA is not considered an option, then there is a danger of the results from the open-label studies being taken at face value, with the strong placebo effect ignored. In health technology assessments, the highest level of evidence of effect is taken from double-blinded trials which specifically attempt to minimise placebo response. If the trials which attempted to control for the placebo effect were excluded then this would inflate the effectiveness of the interventions and potentially result in a recommendation of interventions that are not a cost-effective use of scarce resources. Additionally it could be argued that if OPLA provides comparable benefits to PVP and BKP, then it may be unethical to perform the active interventions which carry a small, but definite, risk of adverse events.

As detailed in the discussion of clinical effectiveness section, there is insufficient evidence to understand the exact nature of the placebo effect. Further research will be needed to determine whether the placebo response can be obtained without resorting to BKP, PVP or OPLA, or whether the observed (potentially psychologically driven) benefits compared with OPM can only be achieved with real or OPLA.

Due to the uncertainty regarding whether OPLA should be included as a comparator (the Johnson and Johnson submission including OPLA, whereas the Medtronic submission did not) where it is indicated that OPLA lies on the cost-effectiveness frontier analyses and where both BKP and PVP have an ICER of >£20,000 per QALY gained, results are additionally presented with OPLA excluded as a comparator.

The decision problem is further complicated by the possibility that BKP and PVP (and potentially OPLA) may have a beneficial effect on mortality. The clinical belief for this is that patients who regain mobility quicker remove fluid from the lungs, regain their appetite and are less prone to infection. Observational data indicate that this is the case. The publications indicating the mortality benefit have been formally critiqued (Appendix 12), and concluded that it is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment. As such, separate analyses are presented where different assumptions regarding the mortality effect have been made.

In addition to the direct use of BKP and PVP, an exploratory analysis was considered by the Assessment Group assuming that all patients are provided with a facet joint injection prior to vertebroplasty as detailed in Wilson et al.[83] Such an intervention is becoming more common in clinical practice according to our clinical advisor.

Stentoplasty was not considered due to the dearth of robust evidence.

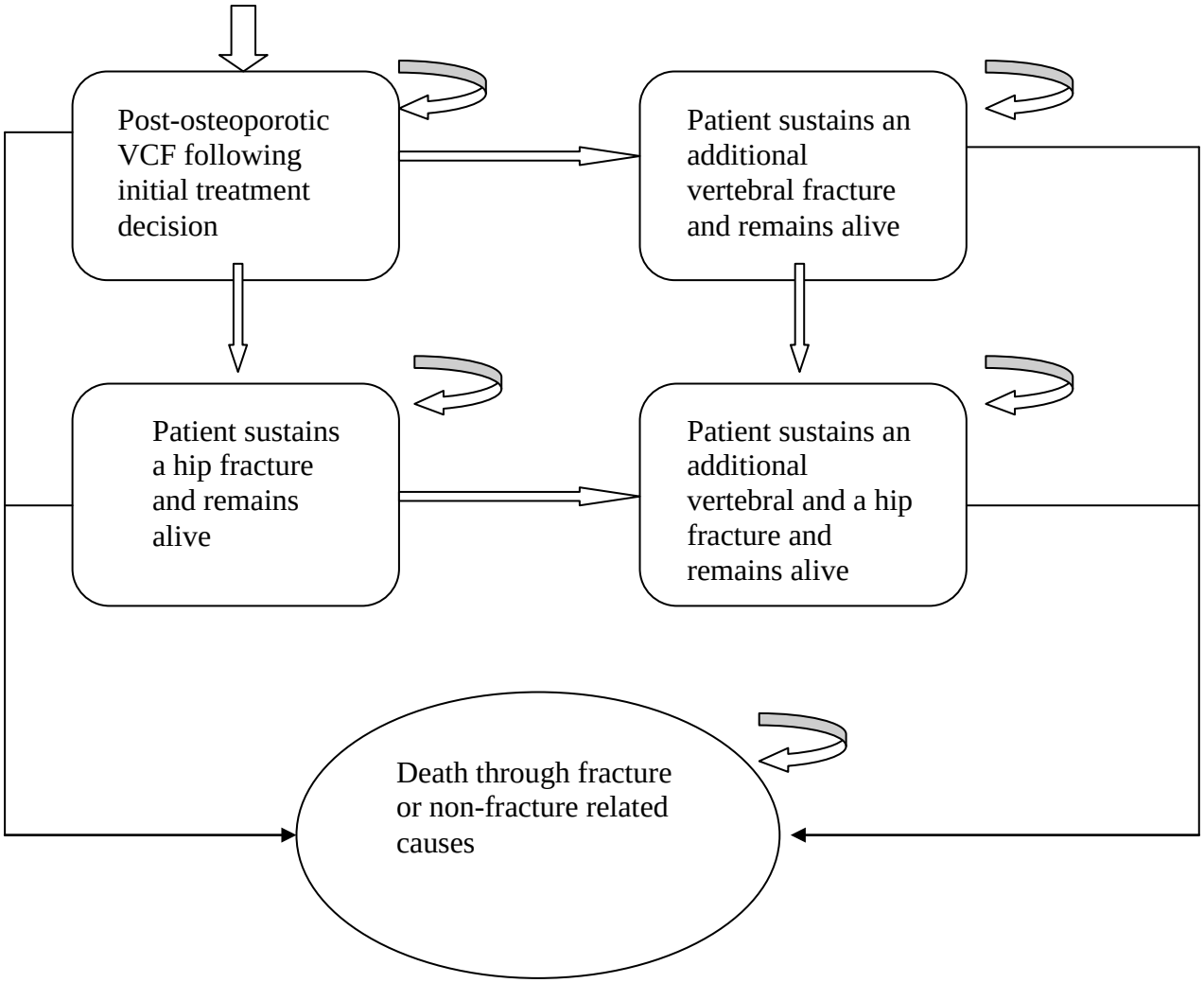
The conceptual model

The conceptual model was constructed to account for two main factors. Firstly the potential difference in EQ-5D (mapped from VAS or taken directly from the trial) within the short term due to the intervention and secondly the need to model differential mortality rates which are dependent on the intervention. As there were potentially different mortality rates it was deemed prudent to also model expensive events related to the osteoporotic VCF to take into consideration the fact that patients who live longer may have other disease related events. Thus, the risks of subsequent hip and vertebral fractures were also modelled.

The model consisted of five health states: post-osteoporotic VCF, for which BKP, PVP, OPLA or OPM has been undertaken (which is the starting state for all patients); a subsequent additional vertebral fracture; a subsequent hip fracture; both a subsequent vertebral and a hip fracture; and death (an absorbing state). For simplicity only one further vertebral fracture and one hip fracture were permitted. The conceptual model is depicted in Figure 21. The model

employed a time horizon of 50 years which was assumed to represent patients' lifetimes and employed 36 monthly time cycles followed by 47 yearly time cycles. The rationale for the different cycle length was that there may be a utility difference between interventions in the initial period following a procedure which was more easily incorporated into monthly time cycles. A life table methodology was employed to take into consideration that all transitions did not take place at the end of the time cycle.[357] Both costs and benefits were discounted at 3.5% per annum.[358] The model did not include the potential disutility associated with anxiety regarding the prospect of future fractures, nor the potential reduction in BMD associated with prolonged bed rest.

Figure 21: Diagram of the conceptual model



The assumed transition probabilities.

Transition probabilities are dependent on the patient age, patient gender, T-Score, the assumed effect of the procedure on mortality, whether bisphosphonates are prescribed and the assumed efficacy of bisphosphonates. A T-Score is defined as the number of standard deviations from the average bone mineral density of healthy young women. For simplicity it was assumed that generic weekly alendronate was the bisphosphonate taken by all patients.

The transition probabilities are detailed in four categories, which represent the health states from which a patient could exit: “Post-osteoporotic VCF following initial treatment decision”; “Patient sustains an additional vertebral fracture and remains alive”; “Patient sustains a hip fracture and remains alive”; and “Patient sustains an additional vertebral and a hip fracture and remains alive”

Transition probabilities from "Post-osteoporotic VCF following initial treatment decision"

- To "Patient sustains an additional vertebral fracture and remains alive"

The transition rates were taken from the values used in Stevenson et al.[359] These values were derived from an exponential fit to population data provided in Singer et al[360] and adjusted to provide a risk for patients with a T-Score of -2.5SD and no previous fracture (Table 20 of Stevenson et al). These values were then multiplied by 1.5 to take into account the additional risks following an initial fracture for patients aged 70 years and over. Further detail on this calibration is provided on p43 and illustrated in Figure 23 of Stevenson et al.[359]

Alternative T-Scores were considered by using the equations provided by Marshall et al[53] which indicate that risk of vertebral fracture increased by 1.8 to the power of the patient's Z-score. The Z-Score is defined as the number of standard deviations from the average bone mineral density of women of the same age as the patient. Assuming that the standard deviation of bone mineral density remains constant as a population ages, the risk of fracture at a T-Score of -3.5SD was therefore assumed to be 1.8 times greater than that of a patient with a T-Score of -2.5SD.

The annual vertebral fracture risks assumed for patients of a given age, and given T-Score on entry to the model are provided in Table 50.

Table 50. The assumed annual risks of vertebral fracture following an initial vertebral fracture based on age and T-Score on entry to the model

	T-Score (SD)			
Age Group (years)	**-2.0**	**-2.5**	**-3.0**	**-3.5**
65-69	0.41%	0.56%	0.74%	1.00%
70-74	0.46%	0.62%	0.83%	1.11%
75-79	0.55%	0.74%	0.99%	1.32%
80-84	0.65%	0.87%	1.17%	1.57%
85-89	0.78%	1.05%	1.41%	1.89%

To take into consideration that a patient's bone density is likely to deteriorate over time, a decrease of 0.255SD per 5-year age group was incorporated in accordance with data from Holt and Khaw[361] presented in Stevenson et al[362] Thus, when a patient became 5 years older the risk of a vertebral fracture increased by $1.8^{0.255}$, which is an increase of 16% compared with a person of the same age, with a T-Score equal to that of the patient's five years

previously. For simplicity it was assumed that women and men with the same T-Score would have the same risks of fracture.

If a patient were assumed to be taking a bisphosphonate, the assumed effect on vertebral fractures was assumed to be a relative risk (RR) of 0.58 (95% CI 0.50 to 0.67) from data reported in Table 27 of Stevenson et al[359] This effect was assumed to last for 5 years, with a linear wane in effect over a 5 year period, so that the RR was 1 after 10 years.

- To "Patient sustains a hip fracture and remains alive"

The transition rates were taken from the values used in Stevenson et al[359] These values were derived from an exponential fit to population data provided in Singer et al[36] and adjusted to provide a risk for patients with a T-Score of -2.5SD and no previous fracture (Table 20 of Stevenson et al). These values were then multiplied by 1.5 to take into account the additional risks following an initial fracture for patients aged 70 years and over. Further detail on this calibration is provided on p43 and illustrated in Figure 23 of Stevenson et al[359]

Alternative T-Scores were considered by using the equations provided by Marshall et al[53] which indicate that risk of vertebral fracture increased by 2.6 to the power of the patient's Z-score. Assuming that the standard deviation of bone mineral density remains constant as a population ages, the risk of fracture at a T-Score of -3.5SD was therefore assumed to be 2.6 times greater than that of a patient with a T-Score of -2.5SD.

The annual risks of hip fracture assumed for patients of a given age, and given T-Score on entry to the model are provided in Table 51.

Table 51. The assumed annual risks of hip fracture following an initial vertebral fracture based on age and T-Score on entry to the model

	T-Score (SD)			
Age Group (years)	**-2.0**	**-2.5**	**-3.0**	**-3.5**
65-69	0.41%	0.56%	0.74%	1.00%
70-74	0.46%	0.62%	0.83%	1.11%
75-79	0.55%	0.74%	0.99%	1.32%
80-84	0.65%	0.87%	1.17%	1.57%
85-89	0.78%	1.05%	1.41%	1.89%

To take into consideration that a patient's bone density is likely to deteriorate over time, a decrease of 0.255SD per 5 year age group was incorporated in accordance with data from Holt and Khaw[361] presented in Stevenson et al.[362] Thus, when a patient became 5 years older the risk of a vertebral fracture increased by $2.6^{0.255}$, which is an increase of 28% compared

with a person of the same age, with a T-Score equal to that of the patient's five years previously. For simplicity it was assumed that women and men with the same T-Score would have the same risks of fracture.

If a patient were taken a bisphosphonate, the assumed effect on hip fractures was assumed to be a RR of 0.72 (95% CI 0.58 to 0.88) from data reported in Table 27 of Stevenson et al[359]. This effect was assumed to last for 5 years, with a linear wane in effect over a 5 year period, so that the RR was 1 after 10 years.

- To "Death through fracture or non-fracture related causes"

The mortality rate associated with hip fracture was taken from Table 21 of Stevenson et al,[359] which reports an estimated 6% of people aged 70-79 years living in the community die from causes related to a hip fracture in the year of fracture. Corresponding figures for patients aged 80-89 years and 90 years and over were 11% and 16% respectively. For simplicity, it was assumed that all patients lived in the community prior to the osteoporotic VCF.

The mortality rate associated with vertebral fracture was taken from a UK study[363] comparing mortality in those with osteoporosis (and no fracture) with mortality in those with osteoporosis and a previous clinically apparent vertebral fracture. The hazard ratio was 4.4 (95% CI 1.85 to 10.6) and was used in the model to inflate the underlying death rate through other causes. The number of years for which a vertebral fracture was assumed to affect mortality rates was user defined with a base case estimate of five years. The effect then linearly dissipated across a user defined period (five years in the base case). When a patient was simulated to have a subsequent vertebral fracture, the model had the facility to allow an increase risk of mortality in the year of subsequent fracture in accordance with data from Jalava.[363] Any effects in subsequent years were not incorporated to limit the number of health states required.

It was assumed that the mortality rate following hip fracture could not be lower than either the mortality rate associated with a vertebral fracture, or lower than that of general mortality in the underlying age and gender matched population. In such circumstances the rate of mortality following hip fracture was increased to equal the higher value.

The underlying death rate through other causes than fracture was taken from http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Interim+Life+Tables (Accessed 30/03/12). For simplicity, it was assumed that all patients would die in their 101st year.

There has been published evidence that mortality may be influenced by initial procedure[215] and further data have been provided in the submission by Medtronic. This is critiqued in Appendix 12. Where BKP, PVP or OPLA were assumed to have positive mortality effects compared with NIM, these were incorporated in the model for a user defined period (set to 5 years in the base case). It was assumed that mortality benefit was not assumed to wane in a linear fashion, but would cease immediately after the user defined period. The relative risks associated with treatment were assumed to apply to the all-cause mortality rates, and the increase associated with vertebral fractures, but not to the value following hip fracture.

Transition probabilities from "Patient sustains an additional vertebral fracture and remains alive"

- To "Patient sustains an additional vertebral and a hip fracture and remains alive"

It was assumed that the risk of hip fracture for patients was independent of whether the patient was simulated to have a subsequent vertebral fracture. Therefore the methodology for calculating the risk of hip fracture was identical to that between the "Post-osteoporotic VCF following initial treatment decision" and the "Patient sustains an additional vertebral fracture and remains alive" states.

- To "Death through fracture or non-fracture related causes"

The methodology for calculating the risk of mortality from fracture and non-fracture causes is identical to that from the "Post-osteoporotic VCF following initial treatment decision" state.

Transition probabilities from "Patient sustains a hip fracture and remains alive"

- To "Patient sustains an additional vertebral and a hip fracture and remains alive"

It was assumed that the risk of vertebral fracture for patients was independent of whether the patient was simulated to have sustained a hip fracture. Therefore the methodology for calculating the risk of hip fracture was identical to that between the "Post-osteoporotic VCF following initial treatment decision" and the "Patient sustains an additional vertebral fracture and remains alive" states.

- To "Death through fracture or non-fracture related causes"

The methodology for calculating the risk of mortality from fracture and non-fracture causes is identical to that from the "Post-osteoporotic VCF following initial treatment decision" state.

Transition probabilities from "Patient sustains an additional vertebral and a hip fracture and remains alive"

- To "Death through fracture or non-fracture related causes"

The methodology for calculating the risk of mortality from fracture and non-fracture causes is identical to that from the "Post-osteoporotic VCF following initial treatment decision" state.

Analyses of the mortality effects associated with the interventions.

A stand-alone critique of the data and methodology used to indicate a mortality benefit associated with PVP and BKP is provided in Appendix 12. This concludes that 'it is possible that there is a causal difference in mortality between patients treated using OPM and OP patients given the size of the effect. Appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. It is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment based on the data presented in the studies included here.'. There is also considerable uncertainty were BKP and PVP assumed to have a mortality benefit, in whether OPLA would also produce a mortality benefit. However, there were no data on this.

Given that there is considerable uncertainty in whether there is a mortality effect it was deemed prudent to explore three scenarios: that BKP had the greatest effect, followed by PVP and then OPM; that BKP and PVP had the same effect which was beneficial compared with OPM; and that BKP, PVP and OPM had the same long-term mortality outcomes. The effectiveness of OPLA was varied in sensitivity analyses. The evidence that was deemed most appropriate was taken from the Cox regression performed on the osteoporotic VCF group who had survived beyond a year that was reported in the Exponent[174] report contained in the Medtronic submission. These values are contained in Table 52, and have been inverted compared with the Exponent report to compare each treatment with OPM, rather than presenting OPM compared with each treatment.

Table 52: The hazard ratios within the three scenarios used to explore the effects of mortality associated with BKP, PVP and OPM

	Hazard Ratio (95% CI)	
Scenario	BKP	PVP
Differential effects	[redacted]	[redacted]
Pooled Effects	[redacted]	[redacted]
No effect	1	1

All values compared to OPM. A lower number indicates that the intervention is associated with a longer life expectancy

Data regarding the effect of OPLA on mortality was not available. For initial analyses it was assumed that the effect was half of that observed for PVP, as this was observed with VAS data (detailed later) which equated to a hazard ratio of [redacted] when a differential effect was assumed, [redacted] when a pooled effect was assumed and 1 when no effect was assumed. The effect of OPLA on mortality was adjusted in sensitivity analyses to acknowledge the arbitrary value used in the initial analyses.

Analyses of the VAS scores associated with the interventions.

Each trial presented results in terms of VAS scores. These are shown graphically, by intervention in Figures 8 to 13. From a visual inspection it appeared plausible that the underlying VAS scores had stabilised at 1 month post-intervention for patients treated with PVP, BKP and sham. However, for OPM the time to stability appeared longer than for the PVP, BKP and OPLA, and was assumed to be 3 months. It was assumed that the VAS scores would be independent of the type of cement used in the procedure.

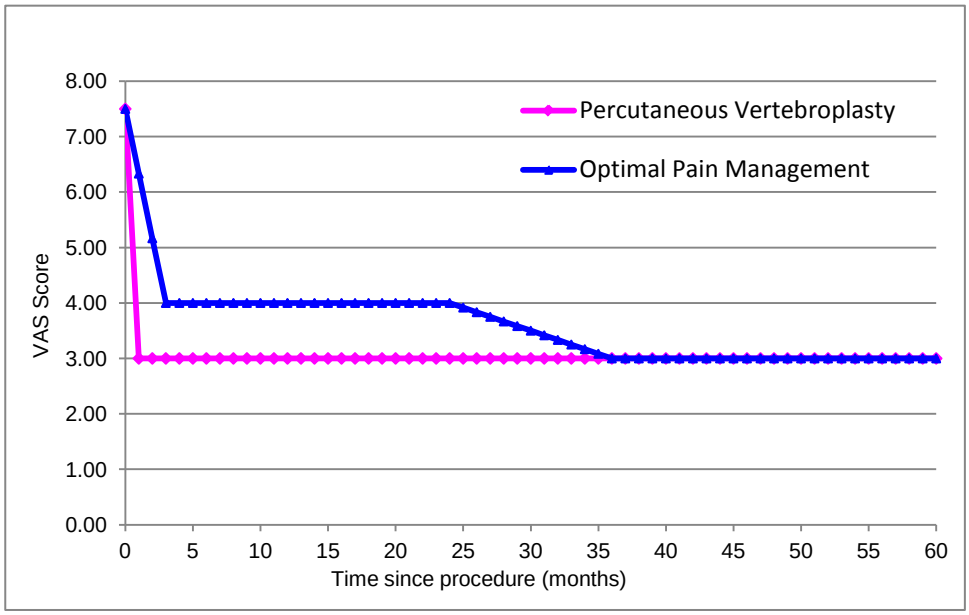
It is assumed that the stable utility following an active intervention remains constant until either the patient moves to another health state, or this value is greater than the underlying population norm value at the patient's age adjusted for the impact of a vertebral fracture (assuming that a vertebral fracture was associated with an ongoing utility multiplier of 0.909 as detailed in Table 24 of Stevenson et al).[359] In the latter circumstance the utility was set equal to the adjusted population value for the given age. The assumed values for the utilities of the population was taken from the mean values reported by Ara et al.[355] The rationale for choosing a constant utility was an analysis of the VAS data for the active interventions which is shown in Figures 8 to 13. The results for each trial are depicted in Appendix 11.

Additionally, it was assumed that eventually the utility would be the same regardless of treatment. On a combination of clinical advice and data from the Farrokhi[147] and FREE[151]

trials, which showed a differential persisting beyond one year, it was assumed that at two years, the utility difference between different treatments would begin to converge in a linear fashion, such that at three years post osteoporotic VCF the VAS scores for all treatments were equal to the VAS score for the treatment generating the greatest patient benefit.

This methodology is illustrated in Figure 22 using the following assumptions: that there are two treatments; (PVP and NIM) the starting VAS score of 7.5, that the stable VAS score for PVP was 3.0 whilst the stable score for NIM was 4.0; and that no further events have happened at 5 years post osteoporotic VCF

Figure 22. An illustrative example of methodology regarding VAS scores post-intervention



Estimation of the initial VAS scores of patients within the studies

The initial VAS data from both arms of each trial were analysed using WinBUGS® to estimate the likely distribution of initial VAS scores for similar patient populations. The CODA output from WinBUGS® was used within the model, however summary statistics are provided for the reader. The mean was a VAS of 7.36, with the 2.5th percentile and the 97.5th percentile values being 6.18 and 8.53 respectively.

Estimation of the stable VAS scores for the interventions.

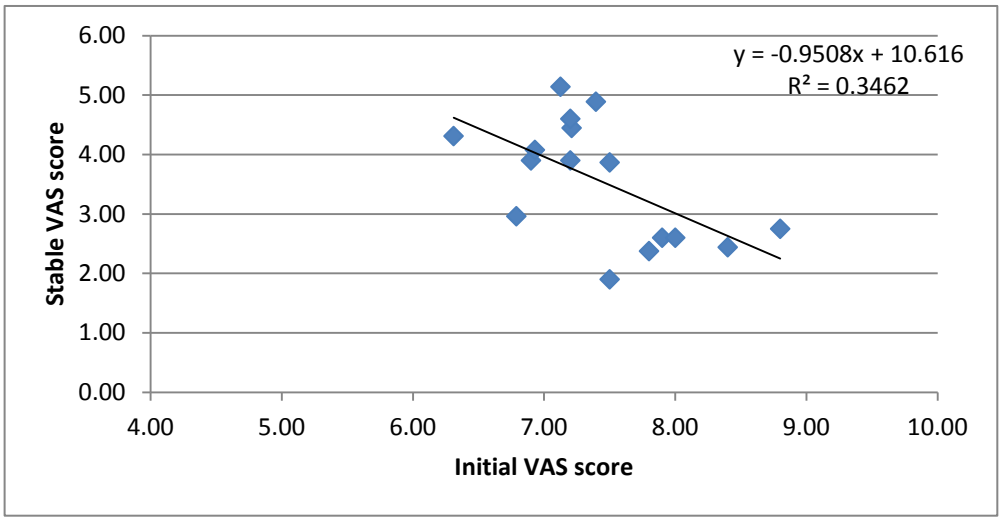
The estimated stable VAS scores, which were assumed to occur at one month post operation for PVP, BKP and OPLA, and at three months post 'treatment' for OPM were calculated in WinBUGS®. A stand-alone report on this process (by Dr Sofia Dias and Professor Tony

Ades) is contained in Appendix 13, with key messages summarised within the main text. Within this process it was assumed that all data was not confounded. The level of crossover is reported in Table 6. In three trials the crossover rate was more than 25% in an arm. These were VERTOS[153], where data was not used to estimate the stable VAS, INVEST[102], where only data at 1 month was used as crossover was prohibited up to this point, and Farrokhi et al[147], in which 26% of patients crossed over from control to PVP; in all other trials the crossover rate was below 25%. The failure to control for crossover is likely to result in a bias against vertebral augmentation, however there was insufficient evidence to allow robust adjustments.

Investigation of the appropriateness of assuming a stable VAS score.

Analyses were undertaken to ascertain if the assumption that the stable VAS score was independent of initial VAS score was appropriate. Figure 23 provides data on the initial VAS and the simple average of VAS scores within the stable period, defined as 1 month and beyond for BKP, PVP and OPLA, and 3 months and beyond for OPM. In order to allow the graphs to be presented the academic-in-confidence VAS data from Buchbinder et al[101] have not been shown. The exclusion of these data did not alter the broad conclusions.

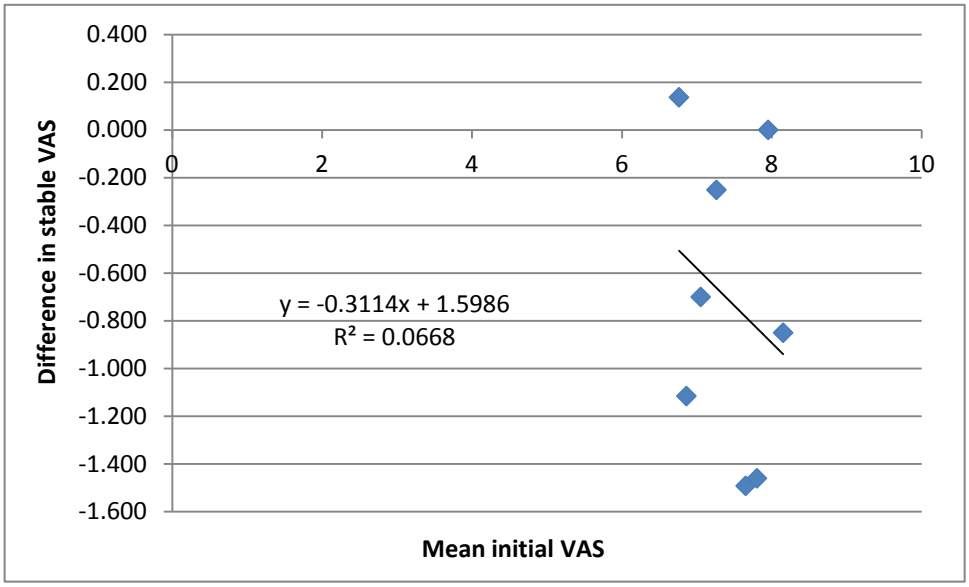
Figure 23. The relationship between initial VAS score and stable VAS score



A moderate relationship was shown between initial VAS score and stable VAS ($r^2 = 0.346$). However, it was seen that the greater the initial VAS score, the lower the stable VAS score; it is unclear whether this is caused by: fractures causing more pain being more responsive to treatment; that there are psychological aspects and that the same pain is rated differently if the preceding pain was worse, or whether the relationship observed is through chance. This conclusion held when analysing the initial VAS score against last VAS score recorded (data not shown)

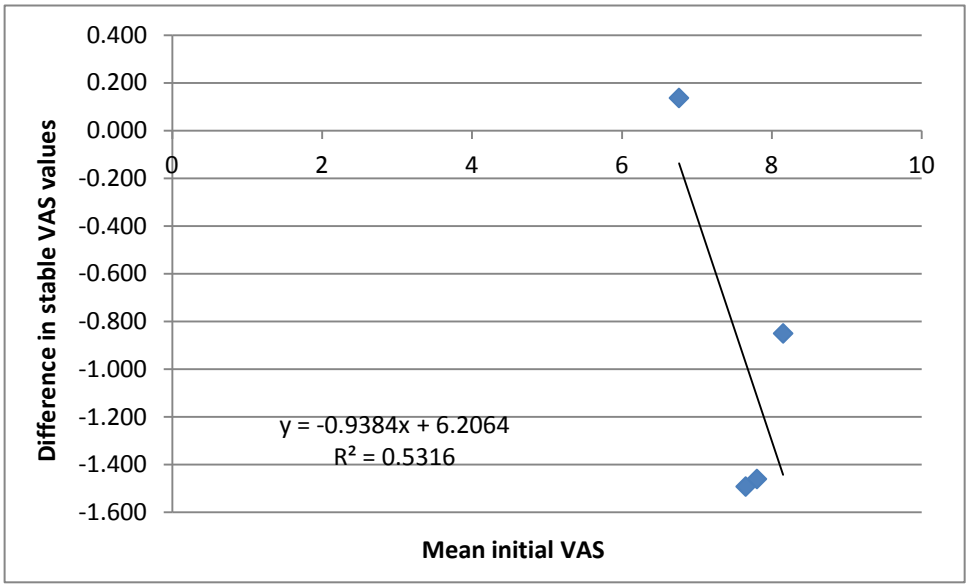
Further analyses were undertaken to ascertain whether the potential relationship between initial VAS score and the stable VAS score could bias the results. An analysis of the difference in the average VAS (the mean of the two arms) at the start of a trial, and the difference in stable VAS scores is presented in Figure 24.

Figure 24. The relationship between the mean initial VAS score and the difference in the stable VAS score



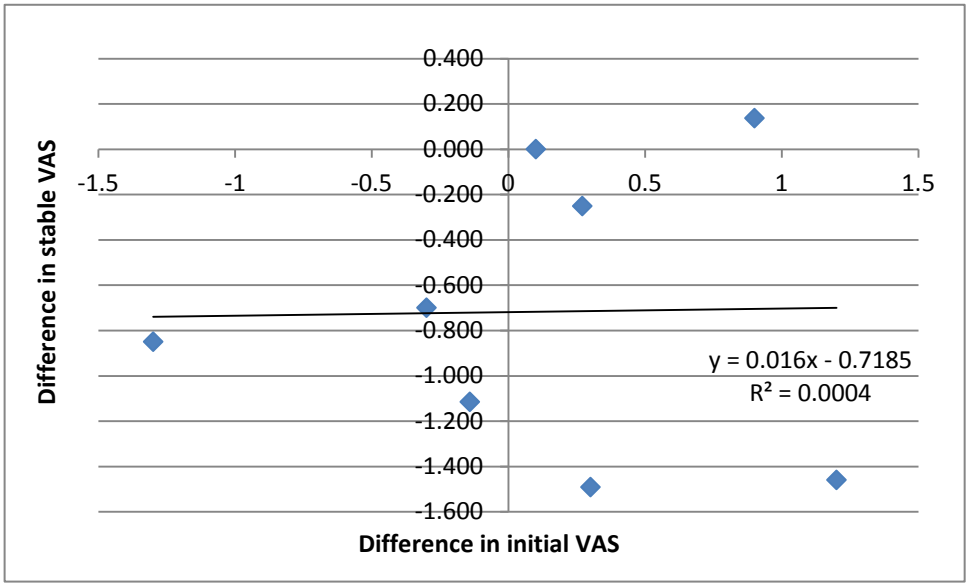
This indicates that there was largely little correlation between the difference in the stable VAS and initial VAS values (r^2 = 0.06). However, this data could be confounded by the different intervention being compared, so a repeat analysis just using the PVP versus OPM trials was conducted (Figure 25). This showed a better fit (r^2 = 0.53) but it is unclear the effect that a smaller number of data points has had on this (by definition a regression of only two data points would have an r^2 value of 1)

Figure 25. The relationship between the mean initial VAS score and the difference in the stable VAS score PVP vs. OPM trials only



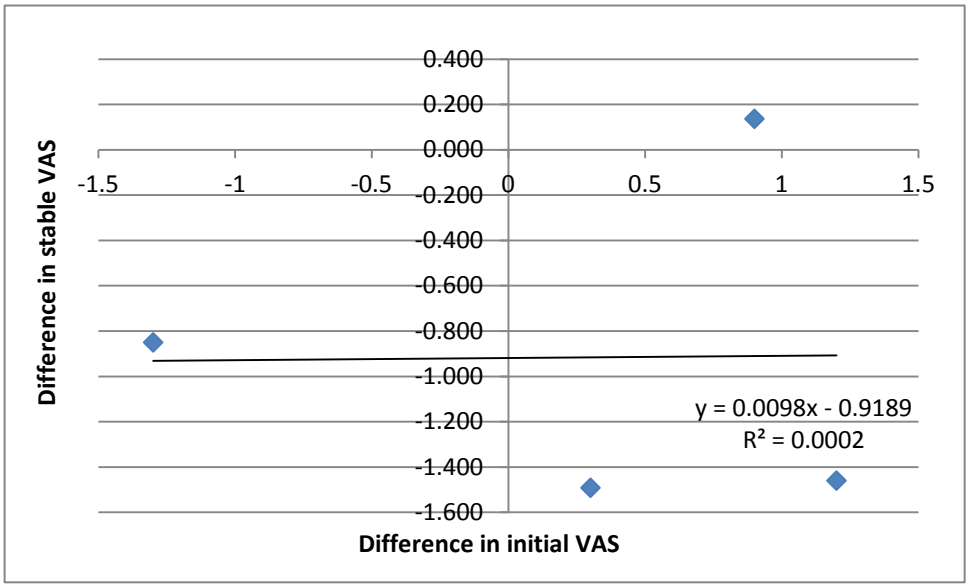
Therefore, an analysis of the difference in the initial VAS values between the arms of the trial and the difference in the stable VAS values was undertaken. This is shown in Figure 26.

Figure 26. The relationship between the difference in the initial VAS score and the difference in the stable VAS score



This indicates that there was largely little correlation between the difference in the stable VAS and initial VAS values ($r^2 < 0.001$). Similar conclusions were drawn when analysing only the PVP vs. OPM trials (Figure 27).

Figure 27. The relationship between the difference in the initial VAS scores and the difference in the stable VAS scores using only the OPM vs. PVP trials.



The authors of this report believe that given the analyses undertaken (Figures 25 to 29) there appears to be little bias introduced by assuming a stable VAS score which is not dependent on the initial VAS score.

Results from the Network Meta-Analyses.

The network of evidence is depicted in Figure 28, with the WinBUGS® output depicted in the form of a caterpillar plot in Figure 29. For a full discussion of the methods used see Appendix 13. Note that there were considered too few trials (four) to undertake meta-regression upon the four treatments.

Figure 28 The network of evidence regarding VAS scores

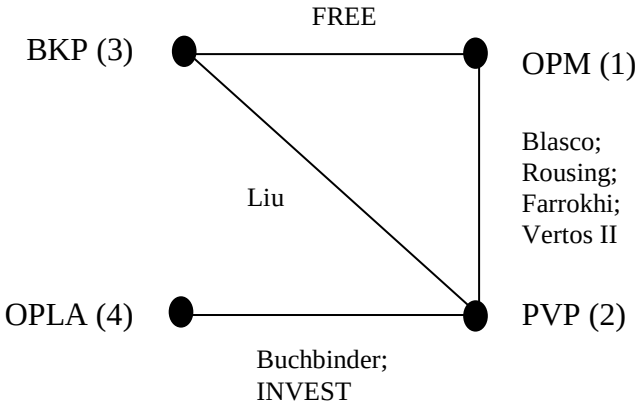
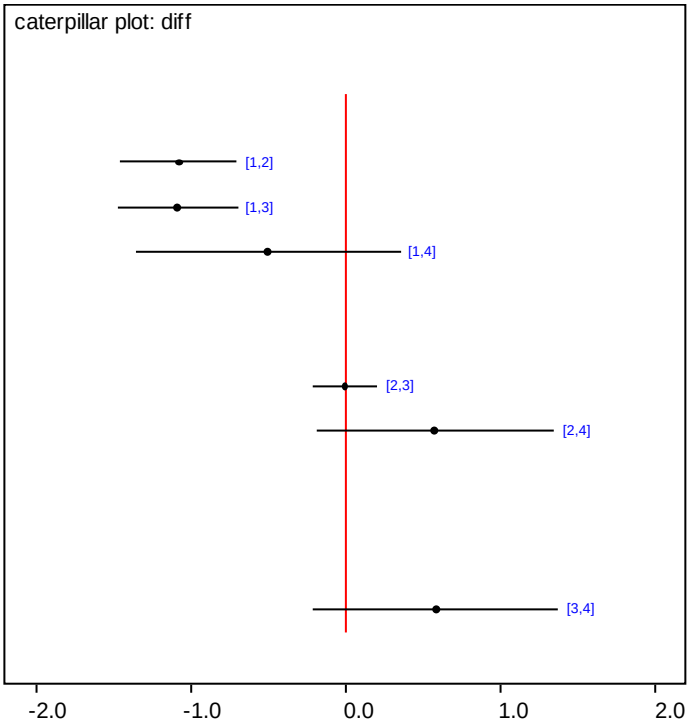


Figure 29 The relative VAS scores



Within Figure 24, 1 denotes OPM, 2 PVP, 3 BKP and 4 OPLA. Values to the left of the no-effect line indicate that the higher numbered intervention produces a lower VAS score.

Thus, a summary of the results estimated by the network meta-analyses of VAS scores is as follows: BKP and PVP appear to be the best treatments with very little difference between them; OPM is significantly worse than both BKP and PVP; there is a possibility that OPLA may produce equivalent results to BKP and PVP or equivalent results to OPM. However, as with standard meta-analyses the quality of the evidence should be appraised, and it is stressed that only two of the trials are of the highest standard (Buchbinder and INVEST). These trials also recorded EQ-5D data which indicate that with respect to change in EQ-5D from baseline that OPLA was equal or marginally inferior to PVP. The implications of this caveat are explored within the results produced by the Assessment Group by undertaking multiple analyses as detailed later.

Mapping Analyses.

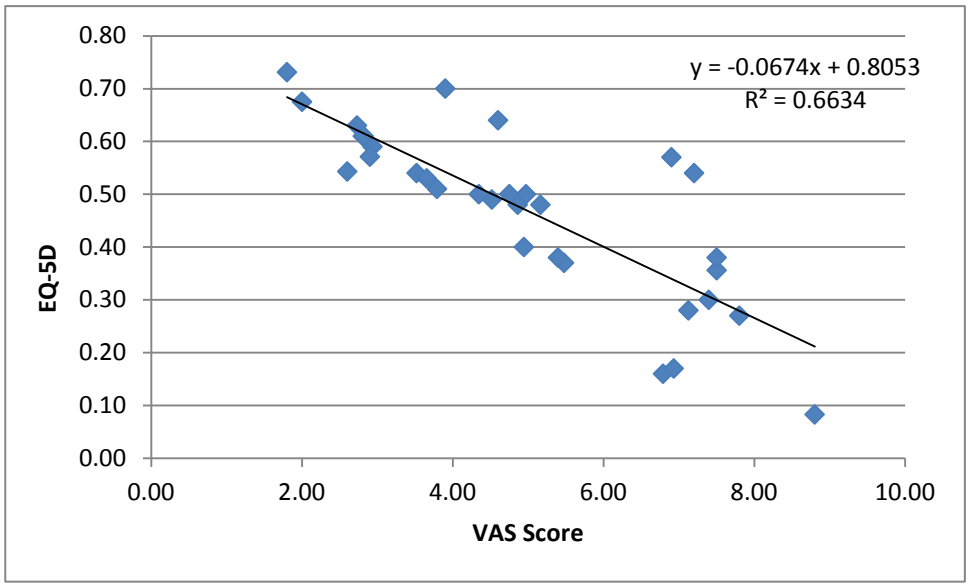
Only a few studies incorporated the EQ-5D, which is the metric recommended in NICE's reference case.[358] To meet the reference case a mapping between an alternate metric and the EQ-5D was required. The initial mapping used VAS scores as all of the studies included incorporated a measure of pain using the VAS score. However, as detailed in the clinical section VAS scores are subjective and may be confounded. Due to this, an analysis of the relationship between RDQ and EQ-5D was also conducted. As the data was provided only at

aggregated level the mapping was undertaken using the mean values for VAS and RDQ scores and EQ-5D. Mapping has the advantage of incorporating data from all studies and thus will not discard data, although will not be as precise as using EQ-5D directly from the trials. Analyses directly using the EQ-5D data reported in the trials have also been conducted and are described later.

Mapping between VAS and EQ-5D.

Data providing both EQ-5D and VAS scores were taken from the FREE study,[151] Buchbinder,[101] INVEST,[102] VERTOS II[17] and Rousing.[185] Data was obtained from the authors of the FREE study.[151] via personal communication. The plot of absolute VAS and absolute EQ-5D is shown in Figure 30, and indicates a relatively good fit, with an r^2 of 0.62. The resultant formula was EQ-5D = 0.8053 – 0.0674 * VAS. The variance on the intercept was 0.00216, the variance on the slope was 0.00008, with a covariance of -0.00038 with these values used in probabilistic sensitivity analyses.

Figure 30 A plot of absolute VAS versus absolute EQ-5D

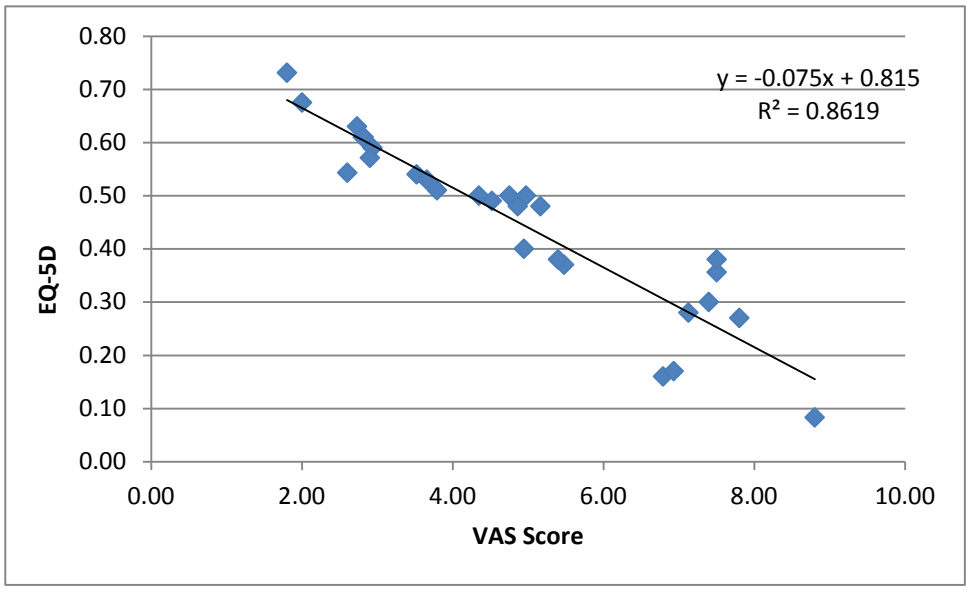


A further analysis was undertaken removing the Rousing[185] and VERTOS II[17] studies as these had used a continuous VAS scale, whereas the remainder of studies had used a numeric rating scale. However, this reduced the explanatory power of the fit (r^2 = 0.50) and the full data set was used, assuming that the continuous VAS scale and the numeric rating scale were interchangeable.

It was possible that the four points considerably above the line were potentially outliers as they all came from the same study (INVEST)[102]. If these data were not included in the mapping, as shown in Figure 31 the fit improved considerably (with an r^2 of 0.86). The

resultant formula was EQ-5D = 0.8392 – 0.0722 * VAS. The variance on the intercept was 0.00095, the variance on the slope was 0.00003, with a covariance of -0.00017 with these values used in probabilistic sensitivity analyses.

Figure 31 A plot of absolute VAS versus absolute EQ-5D excluding INVEST data



Mapping of Roland-Morris Disability Questionnaire (RDQ) onto EQ-5D

Figure 32 and 33 show the relationship between RDQ and EQ-5D dependent on whether INVEST data were included. The r^2 values were 0.55 using the full data and 0.84 excluding the INVEST data.

Figure 32 A plot of absolute RDQ versus absolute EQ-5D

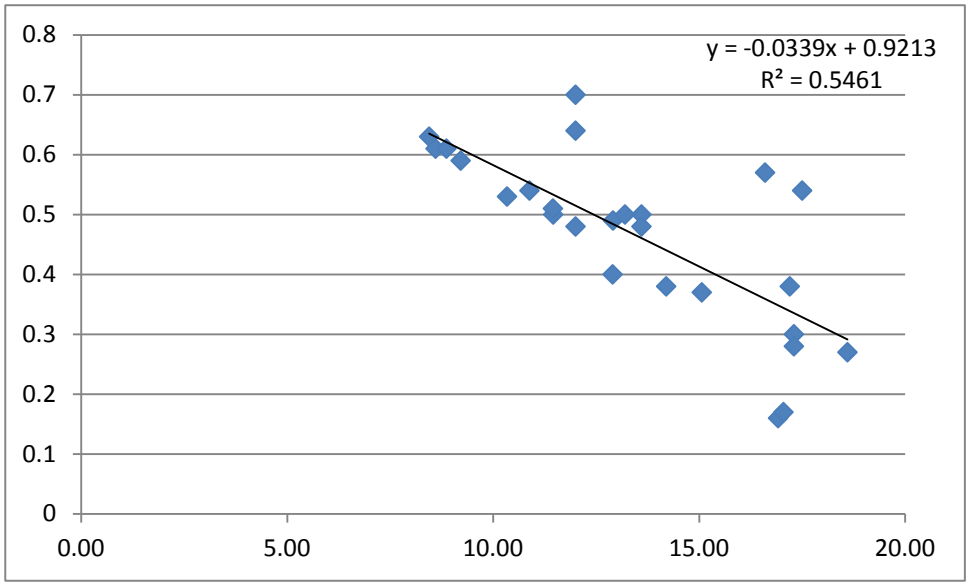
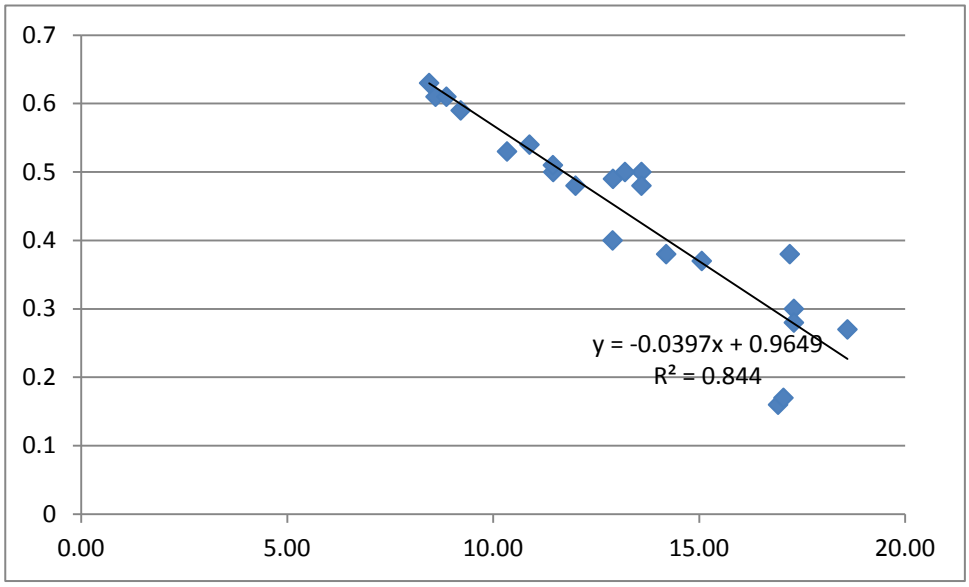


Figure 33 A plot of absolute RDQ versus absolute EQ-5D excluding INVEST data



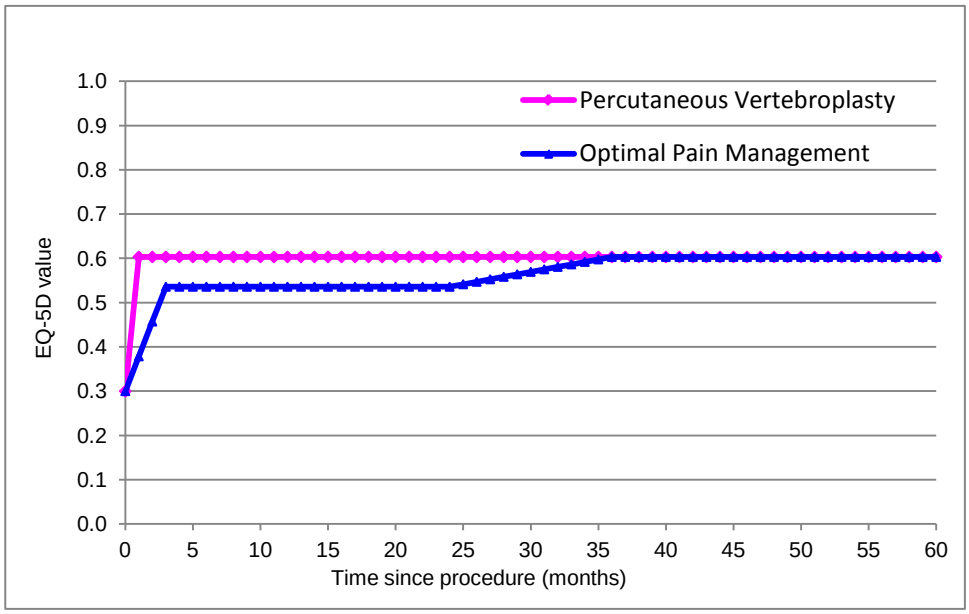
Choice of mapping methodology used.

It was seen that the r^2 values were greater in the VAS and EQ-5D regression than in the RDQ and EQ-5D regression. Given that all studies reported VAS scores, whilst not all studies reported RDQ that it would be preferable to assume a mapping from VAS to EQ-5D rather than RDQ to EQ-5D.

Both statistical relationships (with and without the INVEST[102] data) were used in the modelling to test the robustness of the results to the choice of fit. It is commented that the better fit without the INVEST data does not mean that this trial should be excluded; it could be that the mapping between VAS and EQ-5D does in reality contain noise, and it is noted that the INVEST data was removed after evaluating the initial regression. It is further commented that the mapping has been undertaken assuming that all points from all trials have been given the same weight regardless of study quality.

Thus the VAS score presented in Figure 22 could be converted to an EQ-5D score. For illustrative purposes it is assumed that the deterministic value applies (using all data) and thus EQ-5D = 0.8053 – 0.0674 * VAS, and this is represented in Figure 34. It is assumed that the predicted EQ-5D scores are lower than that of the age and gender mixed population and thus need no adjustment. The area under the curve would then equate to the QALYs accrued post treatment. In this example there would be undiscounted QALY values of 3.00 for PVP and 2.82 for OPM in the initial five year period.

Figure 34 An illustrative example of the conversion of VAS scores to EQ-5D



Utility values

The assumed utility within each health state

The utilities within each state are dependent upon a multitude of factors that are detailed for each health state.

Post-osteoporotic VCF following initial treatment decision

The utility of the patient is assumed to be a function of: patient gender; patient age; the procedure undertaken (NIM, PVP, BKP and OPLA); the time since the procedure; the time at which patients treated with NIM were assumed to have the same utility as patients treated with an active intervention (PVP, BKP and OPLA); the disutility associated with vertebral fractures that occurred greater than one year ago; and the mapping of VAS scores onto the EQ-5D.

Patient sustains an additional vertebral fracture and remains alive

In addition to the factors that influence the utility of a patient in the 'Post-osteoporotic VCF following initial treatment decision state' the disutility associated with a vertebral fracture in the year of occurrence is considered relevant.

In the cycle of the subsequent vertebral fracture, a QALY decrement is automatically applied to account for the associated pain. This QALY decrement is calculated based on the assumed multiplier in the year of the fracture (0.626) and the assumed multiplier in subsequent years (0.909) [359] and the estimated utility score for the patient if no further events had occurred. The patient utility is assumed to be the lower of two values: the population value matched by age

and gender, modified by the prevalent vertebral fracture, and the utility following the initial osteoporotic VCF as depicted in Figure 34 an illustration, were a patient to have an estimated utility of 0.5, then the QALY decrement would be assumed to be 0.206, calculated as 0.5 / 0.909 (the QALY expected in subsequent years following the vertebral fracture) * 0.626 (the QALY expected in the year of the vertebral fracture).

People within this state where the additional vertebral fracture occurred more than one year previously would have the population value matched by age and gender multiplied by 0.909 to take the prevalent vertebral fracture into account.

Patient sustains a hip fracture and remains alive

In addition to the factors that influence the utility of a patient in the 'Post-osteoporotic VCF following initial treatment decision state' the disutilities associated with a hip fracture in the year of occurrence and in subsequent years are considered relevant.

In the cycle of the hip fracture, a QALY decrement is automatically applied to account for the associated pain. This QALY decrement is calculated based on the assumed utility multiplier in the year of the fracture (0.792) and in subsequent years $(0.813)^{359}$ and on the underlying utility scores for a patient of the same age and gender. This calculation uses the same methodology as for patients with a subsequent vertebral fracture who remain alive.

People within this state where the hip fracture occurred more than one year previously would have the population value matched by age and gender multiplied by 0.813.

Patient sustains an additional vertebral and a hip fracture and remains alive

In addition to the factors that influence the utility of a patient in the 'Post-osteoporotic VCF following initial treatment decision state' the disutilities associated with a vertebral fracture in the year of occurrence, with a hip fracture in the year of occurrence and in subsequent years are considered relevant.

This state can be reached from two health states either by sustaining an additional vertebral fracture following a hip fracture or by sustaining a hip fracture following an additional vertebral fracture. In the first route a QALY decrement is applied using the methodology described for "Patient sustains an additional vertebral fracture and remains alive" with the population utility having been adjusted for a prevalent hip fracture. In the second route a QALY decrement is applied using the methodology described for "Patient sustains a hip fracture and remains alive".

People within this state where the most recent fracture occurred more than one year previously would have the underlying population value multiplied by 0.909 (for the prevalent vertebral fracture) and by 0.813 (for the prevalent hip fracture).

Dead

By definition the utility within this health state is zero.

Additional cost and QALY consequences associated with adverse events

The model includes the facility to allow a QALY decrement to be applied to take serious adverse events into consideration. These are calculated crudely based on the likely incidence of serious adverse events and the severity of each event and are subjected to sensitivity analyses. An initial analysis was conducted assuming that there were no cost or QALY implications of adverse events, with a sensitivity analysis conducted assuming that the QALY losses associated with BKP and PVP were 0.02. This value was estimated assuming that the rate of mortality was 1 in 1000 (with an assumed average loss of 10 discounted QALYs) and the rate of morbidity was 1 in 100 (with an assumed average loss of 1 discounted QALYs). When summated this equated to 0.02 discounted QALYs. A threshold analysis is presented to estimate the additional QALY losses at which low viscosity cement would have an equal net benefit to high-viscosity cement assuming a threshold of £20,000 per QALY.

Costs

The assumed costs within each health state

This section focuses on the costs associated with each health state. The values within the health states have largely been taken from Table 25 of Stevenson et al[359] and inflated to 2010/11 prices using the Hospital and Community Health Services inflation indices reported by Curtis et al.[364]

Post-osteoporotic VCF following initial treatment decision

It was assumed that ongoing costs following the initial vertebral fracture would equate to £229 per year.

Patient sustains an additional vertebral fracture and remains alive

The cost of a vertebral fracture was assumed to be £3081, assuming that all fractures occurred in people aged 70 years or greater. This value includes a component for home help. The ongoing costs of £229 per annum associated with vertebral fracture are also continued.

Patient sustains a hip fracture and remains alive

The cost of a hip fracture was assumed to be £7536, assuming that all fractures occurred in people aged 70 years or greater. This value includes a component for home help. For simplicity it was assumed that no patients required nursing home care following a hip fracture, which is acknowledged to underestimate the costs of a hip fracture, although this is not expected to significantly affect the results. The ongoing costs of £229 per annum associated with vertebral fracture are also continued.

Patient sustains an additional vertebral and a hip fracture and remains alive

This state can be reached from two health states either by sustaining an additional vertebral fracture following a hip fracture or by sustaining a hip fracture following an additional vertebral fracture. In the first route the costs are the same as for "Patient sustains an additional vertebral fracture and remains alive"; in the second route the costs are the same as for "Patient sustains a hip fracture and remains alive". The ongoing costs of £229 per annum associated with vertebral fracture are also continued.

Dead

It was assumed that death carried no further cost.

Costs associated with the initial osteoporotic VCF.

The costs associated with the initial osteoporotic VCF have been classified into three categories: the acquisition costs of the interventions, the costs associated with the operation and the costs associated with the length of stay.

The acquisition costs of the interventions

The cost of the CONFIDENCE SPINAL CEMENT SYSTEM^{TM} were taken from the Johnson and Johnson submission, although assuming that 11cc of cement was needed for a 2-level procedure rather than 7cc. This resulted in an average cost of £1546 per operation.

In addition to high-viscosity cement low viscosity cements are also available to purchase at prices that are lower than that of high-viscosity PMMA cement. The list price for such cements were obtained through NICE, and on clinical advice it was estimated that the costs using lower-viscosity cements, incorporating injection kit, needles cement and assorted consumables would be in the region of £660, £720 and £780 for one-, two- and three-level procedures respectively. When weighted for the proportion of operations that are one-, two- and three-level procedures this would equate to an estimated value of £697. However, our clinical expert estimated that 15% of cases are more complex and would require Cortoss®

cement, collation or thicker cement, whilst younger patients would need bone absorbable cement. It was assumed that the added cost of these complex cases would add slightly over £100 to the average cost of an operation resulting in an assumed cost of £800 per low-viscosity cement PVP procedure. Given that the estimate includes a component for using higher viscosity cement, the price used within the analysis could be equated to a strategy where low-viscosity cement is used within the majority of patients, whilst higher-viscosity cements are used in a small proportion where the clinician believes that this is appropriate. Sensitivity analyses were undertaken on these average values.

The list price of BKP (£2600.50 per kit) has been inflated to take into consideration that a proportion of patients will require BKP at more than one level. On clinical advice it was assumed that the percentages reported for PVP were also applicable to BKP, an assumption also stated in the Johnson and Johnson submission. On clinical advice it was assumed that each level would need an additional pack of Kyphon® HV-R® Bone Cement priced at £62 per pack with the remaining instruments being reused. This resulted in the average price per patient increasing to £2639 for BKP. It is noted that this is noticeably less than the £4202 that would be predicted were a new kit required for each level as is implied in the Medtronic BKP brochure. It is commented that our value is significantly higher than that assumed by Medtronic as they did not use the list price but used the average selling price. The NICE Methods guide[358] (section 5.5.2) is clear that 'Analyses based on price reductions for the NHS will only be considered when the reduced prices are transparent and can be consistently available across the NHS, and if the period for which the specified price is available is guaranteed.'. As such only the list price is used.

The cost of OPLA treatment is contentious. The Medtronic submission did not consider OPLA to be a comparator. Johnson and Johnson submission did consider OPLA a comparator but assumed that the cost of this treatment was equal to that associated with PVP. It is uncertain, given the nature of the OPLA the extent of any cost savings compared with vertebroplasty. The impact of the potential cost savings have been evaluated within sensitivity analyses. Table 53 summarises the acquisition costs of the interventions assumed by the assessment group.

Table 53. The acquisition cost of each intervention assumed by the assessment group

Intervention	Assumed cost of intervention
PVP –High viscosity	£1546
PVP – Low viscosity	£800
BKP	£2639
OPLA	To be explicitly considered in a sensitivity analysis – see text
OPM	£0

The costs associated with the operation

The costs of the preliminary phase, the operating phase and the post-operative phase have previously been reported in Ström O et al.[349] and both Johnson and Johnson reported that these prices were inflated. Howevever, there were discrepancies between the two submissions in the values reported (£1479 for Johnson & Johnson and £990 for PVP and £1013 for BKP in the Medtronic submission). Our clinical expert (DW) reviewed the values reported by Johnson and Johnson. Whilst it was deemed there were discrepancies with current UK practice (for example.in the description of the clinician seeing the patient; in the potential overuse of spinal X-Rays and that the operation would most likely take place in an interventional suite rather than an Operating Room) it was concluded that the prices were broadly correct for the preliminary and post-operative phases. Therefore these were used by the Assessment Group. However, for the operating phase the expert was of the opinion that the bottom-up costs provided by Johnson and Johnson were a more realistic estimation than those of Strom et al, and thus these were used, although these were marked as academic in confidence by the manufacturer.

Table 54. The total preliminary, operating and post-operative costs assumed in the Assessment Group model

Phase	Estimated Cost
Preliminary Phase	£540
Operating Phase	£528
Post-operative Phase	£243
Total Cost	£1311

The costs associated with hospitalisation stay

The length of stay following each intervention.

There appears to be considerable uncertainty regarding the lengths of stay associated with each intervention. It is noted that our clinical advisor was surprised by the values presented by both manufacturers (summarised in Table 55), commenting that the majority of interventions (PVP or BKP) are undertaken in Oxford as day procedures and that patients would not be admitted to hospital to have these interventions performed. This is further reinforced by the synthesis of data reported within the pivotal trials which do not indicate a significant length of stay following any of the procedures.

It is commented that the manufacturer of PVP presented data where PVP had the shortest length of stay, whereas the manufacturer of BKP presented data where BKP had the shortest length of stay.

Table 55 The estimated length of stay in days (standard error) assumed in the manufacturers' submissions.

Intervention	Johnson and Johnson	Medtronic
PVP	3.24 (0.49)	6.2
BKP	4.48 (0.89)	5.1
OPM	12.61 (0.27)	9.5

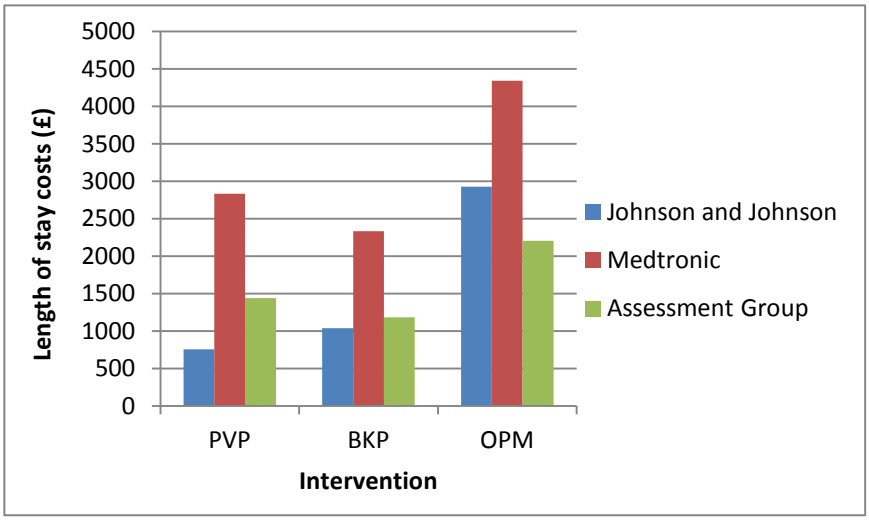
Johnson and Johnson assume that the costs of OPLA are equivalent to PVP

The assumed hospitalisation costs per day.

The manufacturers again present divergent results. Johnson and Johnson assumed a cost of £232 per day based on the Payment by Results national tariff price for an excess bed day associated with VP/KP and NIM healthcare resource group (HRG) codes (HRGs HC04C, HC05C and HD36C).[353] Medtronic assumed a cost of £457 per day for hospitalisation, citing NHS Reference costs 2009/10/11. These values are summarised in Table 56 and Table 57.

The different estimates of hospitalisation costs are depicted in Figure 35. The values assumed by the Assessment Group (which are detailed below) are also shown for information.

Figure 35 The different assumed hospitalisation costs within the manufacturers submission and Assessment Group base case



In the analyses undertaken by the Assessment Group, it was decided that the length of stay data from Medtronic was most appropriate as it used the standard HES data source, and the cost values presented by Johnson and Johnson (£232 per day) were most appropriate. Whilst the £232 per day value was acknowledged to likely underestimate the total costs in each arm, it was deemed more likely to accurately assess the incremental difference between strategies, which would relate to the latter part of hospital stays. As it is the incremental differences rather than the total values that are used in the cost-effectiveness calculations this approach was assumed reasonable. In addition, sensitivity analyses were performed using both the Johnson and Johnson values and the Medtronic values. In addition, an exploratory analysis looking at the impact on the results were the length of stay data assumed equal for all interventions, which was operationally achieved by setting the assumed length of stay to zero for all interventions. The values assumed by the Assessment Group are shown in Table 57. It was assumed that the outcomes associated with PVP were also applicable to OPLA. It was assumed that the ratio of standard error to mean associated with the Johnson and Johnson length of stay data was applicable for the Medtronic data.

The Assessment Group values were consistently lower than those of Medtronic due to the lower cost per bed day. Compared with Johnson and Johnson both PVP and BKP were more expensive due to the longer assumed hospital stay, although the costs associated with OPM were lower.

Our clinical advisor was extremely surprised by the length of stay information, commenting that in the more than two thousand vertebral augmentation procedures he has undertaken the average length of stay would be less than 6 hours. To incorporate this information which may

represent current practice better than the bundled HES data, sensitivity analyses would be performed assuming that there was no cost difference in length of stay, which was achieved operationally by setting the cost per bed day to £0. Unfortunately there is a paucity of reported data regarding length of stays within the trial to ensure that the cost and clinical data align. The lack of these data may indicate that the length of stay was briefer than suggested by HES data.

Table 56. The base case estimated length of stay assumed by the Assessment Group model.

Intervention	**Length of Stay (days ((standard error))**	**Cost per day**	**Total Cost per hospital stay**
PVP	6.2 (0.94)	£232	£1438
BKP	5.1 (1.01)	£232	£1183
OPM	9.5 (0.20)	£232	£2204

The values for OPLA were assumed identical to PVP

Comparison of the model structures and population.

Table 57 provides a summary of the comparison of mathematical models structure developed by the Assessment Group, Johnson and Johnson and Medtronic.

Table 57 A comparison of the mathematical models structures developed by the Assessment Group, Johnson and Johnson and Medtronic.

	Assessment Group	**Johnson and Johnson**	**Medtronic**
Model Structure	State Transition Model	Area Under Curve Simulation	State Transition Model
Base case Time Horizon	Lifetime (maximum 101 years)	1 Year	Lifetime (maximum 101 years)
OPLA included as a comparator	Yes	Yes (but assumed dominated)	No
Consideration of differential mortality effects related to treatment	Yes	No	Yes
Network Meta-analysis undertaken to estimate VAS scores	Yes, with an assumption of stable VAS independent of initial VAS	Yes, but with no interpolation or extrapolation and the Blasco trial was published after their search	Discussed but not considered appropriate
Analyses using direct EQ-5D data	Yes	Yes	Yes
Consideration that people with vertebral fracture may have a poorer survival prognosis	Yes (limited to five years) UK data used[363] and assumed independent of age and time since fracture	No	Yes (limited to five years). Age and time since fracture dependent data derived from Sweden used.
Consideration of subsequent vertebral fractures	Yes (limited to one additional). UK fracture rates used	No	Yes. Fracture rates calculated using hip fracture data and hip: vertebral fracture ration seen in Sweden
Consideration of increased mortality after subsequent vertebral fractures	Yes (limited to one year). Data as above	No	Yes (limited to five years). Data as above
Consideration of subsequent hip fractures and associated mortality	Yes (limited to one additional)	No	No
Consideration of serious adverse events related to vertebral augmentation	Yes (in sensitivity analyses)	Assumed None	Assumed None

The aggregated costs associated with the osteoporotic VCF

A summary table presenting the values used by Assessment Group' Johnson and Johnson and Medtronic is given in Table 58.

Table 58 The base case aggregated costs assumed by the manufacturers and the assessment group.

	Assessment Group	**Johnson and Johnson**	**Medtronic**
Acquisition cost of PVP using low viscosity cement where possible	£800	-	-
Acquisition cost of PVP using high viscosity cement	£1546	£1472	£1193
Acquisition cost of BKP	£2639	£2842	£1996
Acquisition cost of OPLA	Evaluated in sensitivity analyses	£1472	-
Operation Costs	£1311	£1479	£990 (PVP) £1013 (BKP)
Hospital stay costs - PVP	Evaluated in scenario analyses, using both manufacturers' values and assuming zero costs	£752	£2833
Hospital stay costs - BKP		£1039	£2331
Hospital stay costs - OPLA		£752	-
Hospital stay costs - OPM		£2926	£4342
Total costs - PVP	Dependent on the scenario analyses being conducted	£3702	£5804
Total costs - BKP		£5360	£5527
Total costs - OPLA		£3702	-
Total costs - OPM		£2926	£4828

The Medtronic submission also included an additional cost of £486 for treating the initial fracture. This value was applicable to all comparators.

Comparison of the results produced by the Assessment Group model when using (largely) the same data as each manufacturer.

In order to assess the level of agreement between the model structures the Assessment Group model was populated so it resembled, as closely as could be achieved relatively easily, each of the deterministic base case models submitted by the manufacturers. This repopulation did not extend to importing: the vertebral fracture rates; the underlying all cause mortality rates; and the underlying population utility assumed by Medtronic,

The results when the assessment group model was populated with Johnson and Johnson data are provided in Table 59. The results when the assessment group model was populated with Medtronic data are provided in Table 60

Table 59 A comparison of the results produced by the Assessment Group model when using Johnson and Johnson base case data

	Treatment	Cost (£)	QALYs	Δ Cost (£)	Δ QALY	ICER (£)
Johnson and Johnson	OPM	£2926	0.507			
	PVP	£3702	0.684	£777	0.177	£4392
	BKP	£5113	0.656	£1410	-0.027	Dominated
Assessment Group	OPM	£2926	0.509			
	PVP	£3702	0.683	£777	0.173	£4480
	BKP	£5113	0.658	£1410	-0.025	Dominated

It is seen that the results are very close, with the discrepancy arising due to the number of days assumed in a year (365.25 in the Assessment Group model and 365 in the Johnson and Johnson model) and the way these interact with the monthly EQ-5D scores which fluctuate across time.

Table 60 A comparison of the results produced by the Assessment Group model when largely using Medtronic base case data

	Treatment	Cost (£)	QALYs	Δ Cost (£)	Δ QALY	ICER (£)
Medtronic	OPM	5394	4.976			
	PVP	6112	5.325	718	0.35	2053
	BKP	6403	5.441	291	0.12	2510
Assessment Group	OPM	6995	6.047			
	PVP	7800	6.474	805	0.43	2057
	BKP	8179	6.708	379	0.23	2508

It is seen that whilst the costs and QALYs predicted in the Assessment Group model are both higher than in the Medtronic model the incremental values are similar and the ICERs very similar, indicating that the differences are unlikely to affect the conclusions. The Assessment Group believe that the discrepancy is caused due to the difference in both the fracture rates assumed (the Assessment Group use vertebral fracture data from Scotland whereas Medtronic

estimate rates based on UK hip fracture data and the Swedish ratio of hip: vertebral fractures) different risks of mortality following subsequent vertebral fracture and the assumed duration of this risk (one year in the Assessment Group model; five years in the Medtronic model.)

The authors concluded that given the results presented in Tables 59 and 60 the programming of the conceptual models into modelling packages was unlikely to be a key driver of the ICER, in comparison to the assumption made regarding the presence of a mortality benefit (which was the cause for the different conclusions in the Johnson and Johnson and the Medtronic model, on the assumed utilities associated with each intervention and whether OPLA should be included as a comparator, and at what cost if so. Accordingly, it was deemed that the Assessment Group model that had produced ICERs very similar to the manufacturers' models when populated with similar data was of sufficient quality to use in the calculation of all forthcoming results.

Methodology for estimating the scenarios to run

There are a large number of potential structural uncertainties that could be evaluated. In order to restrict the quantity of data presented the following methodology was used.

1) Two foundation analyses were established, which represented two of many plausible scenarios. The assumptions and data used in the foundation analyses are detailed below with the difference between the analyses being that one assumed a mortality benefit associated with BKP, PVP and OPLA, whilst one did not. The deterministic results from the foundation analysis were calculated.
2) Univariate sensitivity analyses were conducted varying structural or parameter values. The effect of the change on the net monetary benefit (NMB) (assuming a willingness to pay of £20,000 per QALY) of the foundation analysis was evaluated. The majority of analyses were undertaken on the model assuming no mortality difference, as where a mortality benefit was assumed the impact of this assumption was far bigger than that of the variables altered.
3) If the change in the NMB were deemed minimal by the authors then the structural uncertainty or parameter was not considered a key driver and would not require a separate scenario analysis in the full analyses. An exception to this rule was allowed if the authors believed that the variable could be important due to an interaction with another variable in multi-variate analyses.
4) Those parameters which had a large impact in univariate sensitivity analyses or were believed could have an impact in multivariate analyses were used to derive the scenario analyses and sensitivity analyses undertaken by the Assessment Group.

Net Monetary Benefit has been used for these exploratory analyses as it is relatively simple to assess which intervention is the most cost-effective (denoted by the intervention with the largest value) and also whether the ICER is below £20,000 per QALY gained (the NMB value is greater than £0) or greater than £20,000 per QALY (the NMB value is lower than £0).

The assumptions used in the foundation analyses

The following are the assumptions and parameters used in the foundation analyses. In the analyses conducted without an assumed mortality benefit, the assumed duration of treatment-related mortality benefit was set to zero. The values in parentheses indicate the values tested in univariate sensitivity analyses.

- Patient Age: 70 years (60, 80)
- Gender: Female (male)
- T-Score: -3SD (-2.5, -3.5)
- Length of bisphosphonate use: 5 years (0)
- Fall time associated with bisphosphonates: 5 years (0)
- The assumed duration of a treatment-related mortality benefit: 5 years (0)
- The assumed duration of the RR of mortality following a vertebral fracture: 5 years (0)
- The assumed wane time associated with the RR of mortality following a vertebral fracture: 5 years (0)
- Include an added risk of mortality in year of subsequent vertebral fracture: True (false)
- Costs associated with hospital stay: Assessment Group (Johnson &Johnson / Medtronic; 0)
- Cost of PVP: Low viscosity cement £800 (High viscosity cement £1546)
- Discount rate costs 3.5% (0, 6)
- Discount rate benefits 3.5% (0, 6)
- QALY loss associated with PVP and BKP = 0 (0.02)
- Hazard Ratio on general mortality for BKP and PVP: ([redacted]) ([redacted])
- Mortality effect of OPLA: Half that of PVP (no effect, equal to PVP)
- The regression mapping VAS to EQ-5D (using all data, excluding INVEST data)
- Assumed point at which VAS scores converge: 24 months (12 months)
- Cost of OPLA: Equal to PVP (20%, 40%, 60% and 80% of PVP)

The following parameters were assumed fixed: costs and utility losses associated with fractures; the acquisition costs of BKP; the distributions of efficacy associated with alendronate; the costs of alendronate; the distribution on the risk of mortality following vertebral fracture; the mortality rate following hip fracture; and the general mortality rate.

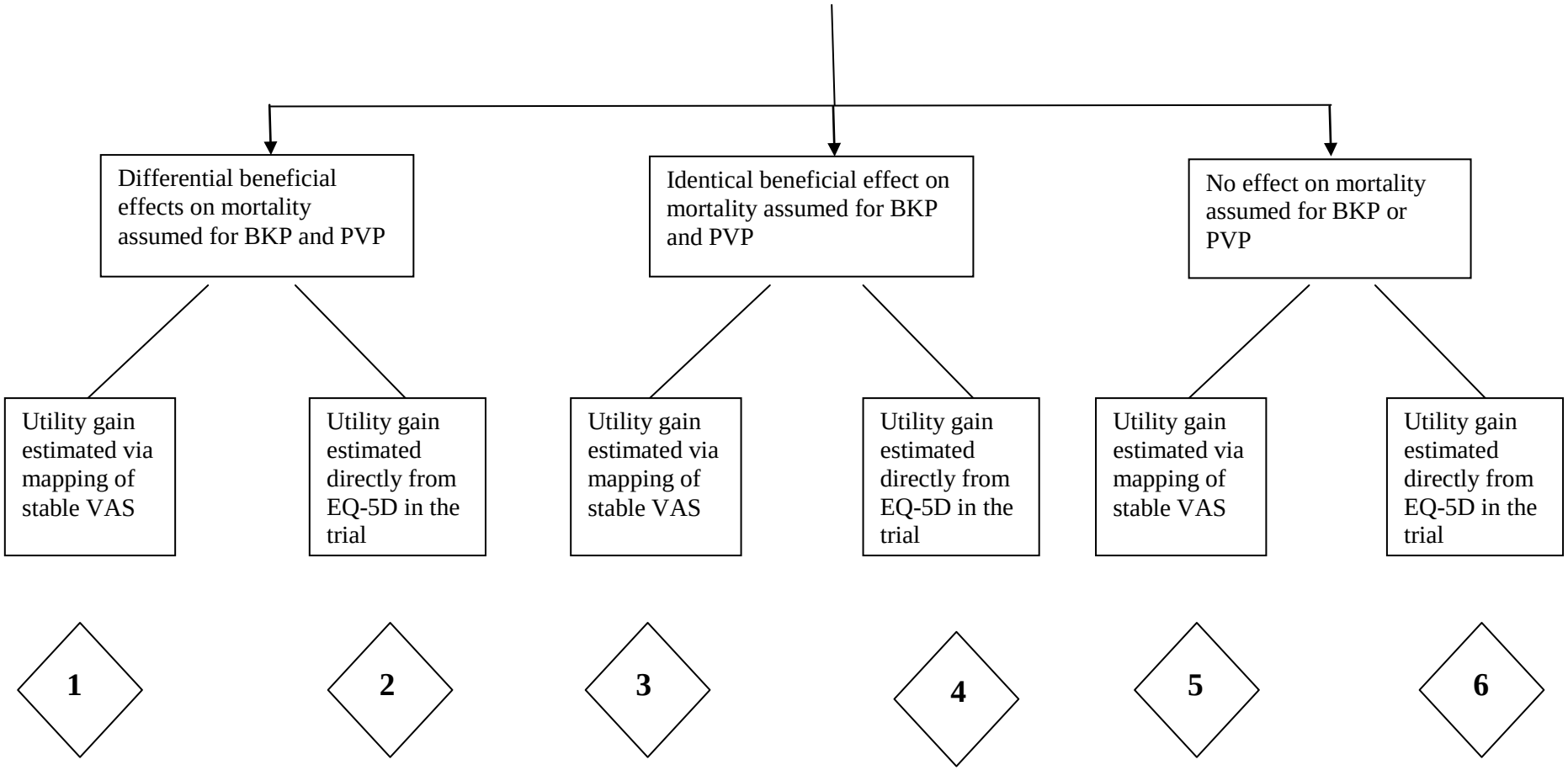
The analyses conducted.

Having undertaken the analyses determining the variables to which the model was sensitive it was clear that the assumption regarding the mortality benefits of the active interventions was crucial to the cost-effectiveness ratios. As such, it was deemed sensible to present the results split into three categories based on the underlying assumption: a differential effect assumed for BKP and PVP with both better than OPM; a pooled analyses where the effects of BKP and PVP were assumed identical with both better than OPM; and where no mortality benefit of BKP or PVP was assumed. The effect of OPLA was varied in sensitivity analyses for the first two categories and assumed equal to OPM in the third.

Furthermore, there was known to be a difference in results based on whether the EQ-5D was taken directly from the RCTs for the four trial reporting such values (INVEST[152], FREE[151], Buchbinder[150], and Rousing[185]) or whether the mapping of stable VAS scores from the network meta-analyses to EQ-5D was the preferred method. Analyses of the change in EQ-5D data showed that the Rousing trial was unable to be used meaningfully due to the imbalance in EQ-5D at the start of the trial between the PVP and control arms. Thus analyses were conducted on just the Buchbinder, Free and INVEST studies.

As such, there were deemed six plausible scenarios that are depicted in Figure 36. The results for these were calculated using both deterministic and probabilistic methods. For each of the scenarios sensitivity analyses were undertaken exploring the effects of changing structural assumptions and parameter values that were deemed important in the exploratory analyses.

Figure 36 Derivation of the Assessment Group's six scenarios.



Sensitivity Analyses will be undertaken for each of the six scenarios.

Scenarios 2, 4 and 6 will also be subdivided into the results from Buchbinder, the FREE trial, and INVEST

The Assessment Group's results

Foundation Analyses

Using the foundation analyses the costs and QALYs associated with each intervention were as depicted in Table 61 and Table 62. An intervention being extendedly dominated indicates that a combination of two other interventions can provide the same health gain at a lower cost. Interventions which are neither dominated nor extendedly dominated form the cost-effectiveness acceptability frontier.

It is stressed that these results do not represent a base case, but one of a number of plausible scenarios.

Table 61 The results of the Assessment Group's foundation analysis assuming no mortality benefit for BKP, PVP or OPLA.

Procedure	Costs	Benefits	ICER (Cost per QALY gained) ¤	NMB†
OPM	£5459	4.74	-	-
OPLA	£6804	4.83	Extendedly dominated	£413
PVP	£6804	4.91	£7802	£2104
BKP	£8388	4.91	Dominated	£514

¤ Compared with the next least effective point on the cost-effectiveness acceptability frontier.

†Compared with OPM at a willingness to pay of £20,000 per QALY gained

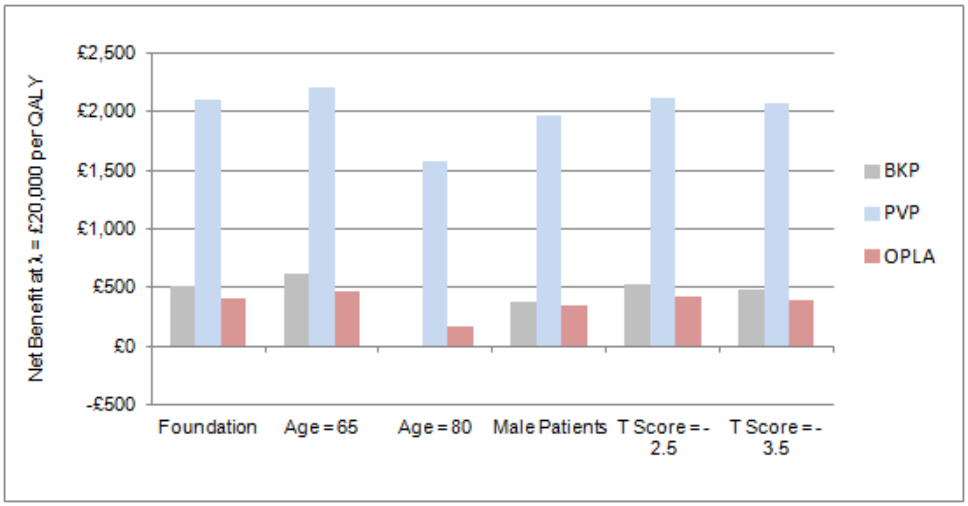
Table 62 The results of the Assessment Group's foundation analysis assuming a relative risk of mortality ([redacted] for BKP, [redacted] for PVP and [redacted] for OPLA).

Procedure	Costs	Benefits	ICER (Cost per QALY gained) ¤	NMB†
OPM	£5459	4.74	-	-
OPLA	£6850	4.89	Extendedly dominated	£1622
PVP	£6897	5.04	£4802	£4550
BKP	£8651	5.27	£7488	£7480
¤ Compared with the next least effective point on the cost-effectiveness acceptability frontier. †Compared with OPM at a willingness to pay of £20,000 per QALY gained				

The results from the Univariate analyses have been grouped into seven categories, the first six of which were tested assuming no mortality benefit. These are illustrated in Figures 37 to 43

- Those that are related to a patient's characteristics (age, gender and T-Score)

Figure 37 Univariate analyses regarding patients' characteristics.

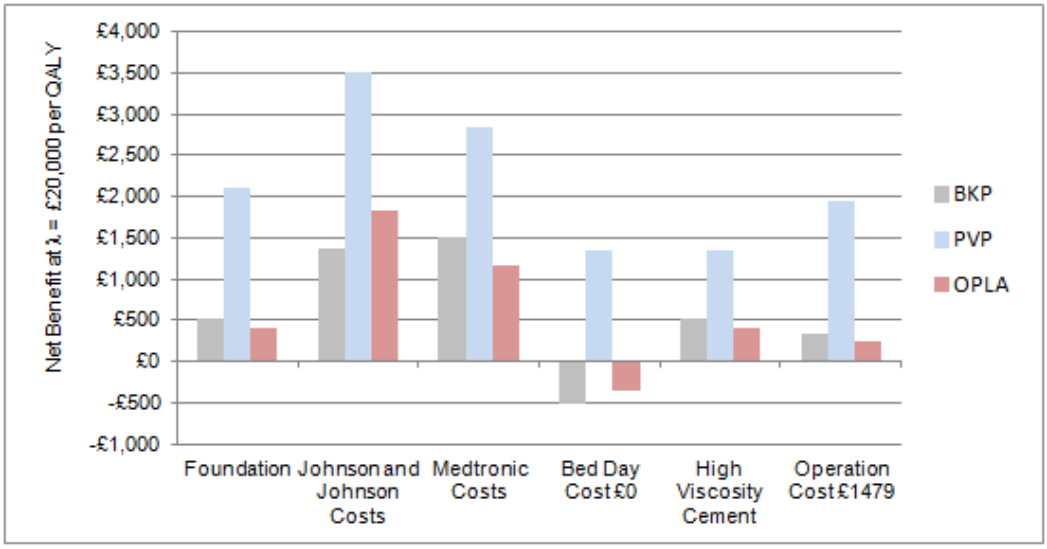


In this univariate analysis, the assumed age of the woman was altered to 65 and 80 years, the gender of the patient was assumed to be male, and the T-Score was altered to -2.5SD and -3.5SD. The authors did not deem that any of these parameters made a noticeable difference to

the conclusions of the foundation analyses, and these variables remained constant in the full analyses.

- Those affecting costs of hospitalisation costs, operation cost and the price of cement.

Figure 38 Univariate analyses regarding hospitalisation costs, operation cost and cement costs.

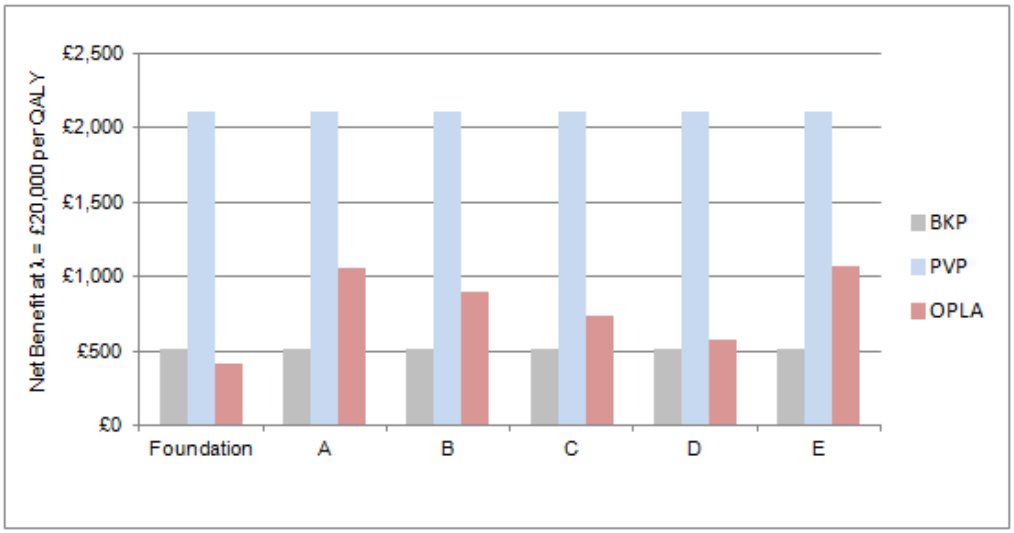


In this univariate analysis the costs of hospital stay was altered to the costs proposed by Johnson and Johnson and Medtronic, the hospitalisation costs were set equal amongst interventions (by reducing bed day cost to £0), the cost of high viscosity cement was used and the operation cost was set to £1479

The authors did not deem that any of these parameters made a noticeable difference to the conclusions of the foundation analyses. However, as the length of stay data are contentious, the analysis using a cost per bed day of £0, which sets the costs equal amongst interventions would be retained. The QALY threshold gain at which high viscosity cement is more cost-effective than low viscosity cement will additionally be calculated.

- Those associated with the costs of equipment required for OPLA and the cost of the procedure when using OPLA.

Figure 39 Univariate analyses regarding the costs of equipment required for OPLA and the cost of the procedure when undertaking OPLA.



Scenario Legend

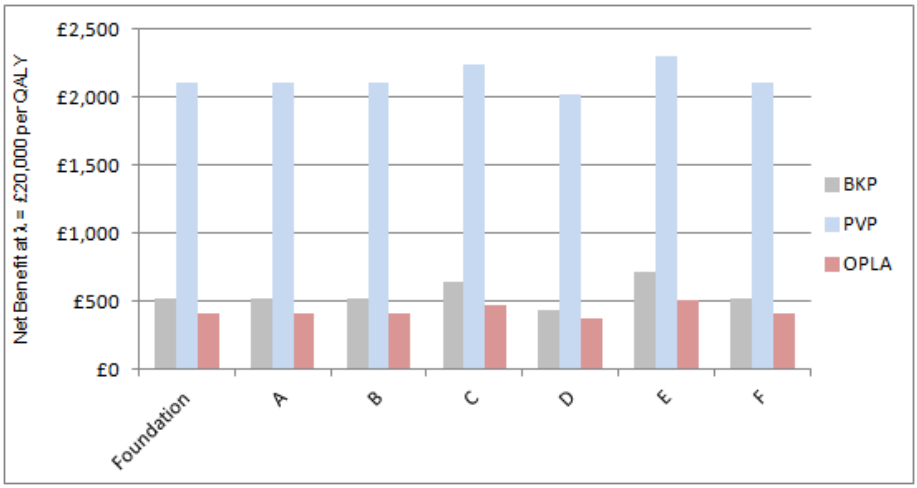
A: Cost of OPLA equipment set to 20% of the cost of PVP equipment
B: Cost of OPLA equipment set to 40% of the cost of PVP equipment
C: Cost of OPLA equipment set to 60% of the cost of PVP equipment
D: Cost of OPLA equipment set to 80% of the cost of PVP equipment
E: Cost of OPLA procedure set to 50% of PVP procedure

In this univariate analysis the costs of PVP equipment and the cost of the OPLA procedure were altered.

The analyses presented did not alter whether OPLA was adjudged to be the most cost-effective intervention. However, because these analyses were purely univariate the effects of a change in both the cost of OPLA equipment and the cost of the OPLA procedure were not calculated. It is plausible that multiple changes would affect the conclusions and these parameters were changed in the full analyses.

- Those associated with the discount rate and bisphosphonate use

Figure 40 Univariate analyses regarding discount rates, bisphosphonate usage and bisphosphonate wane period.



Scenario Legend

A: Discount rate for future costs set to 0% per annum

B: Discount rate for future costs set to 6% per annum

C: Discount rate for future benefits set to 0% per annum

D: Discount rate for future benefits set to 6% per annum

E: An assumption that no woman was taking bisphosphonates

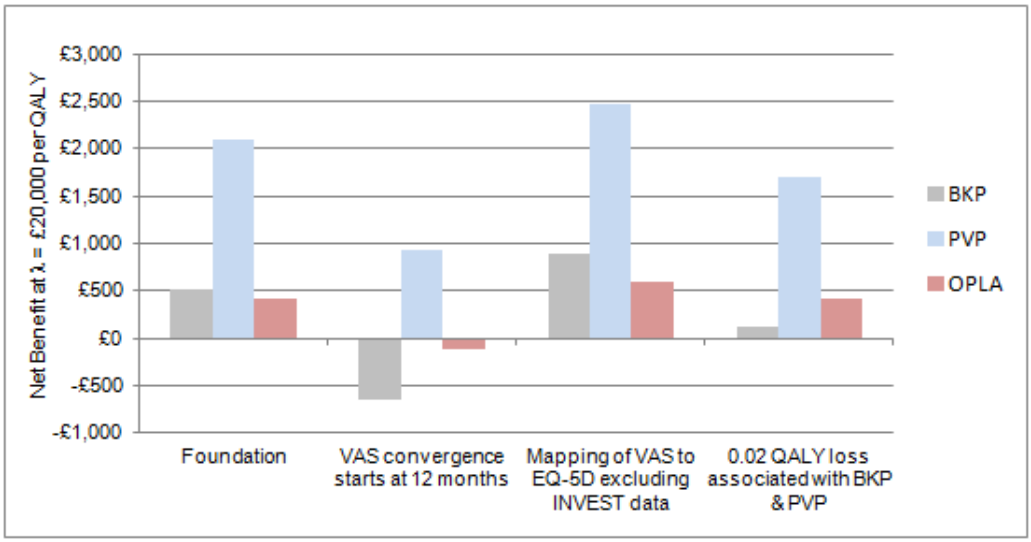
F: The wane period following bisphosphanate treatment set to zero years

In this sensitivity analyses the discount rates for both costs and benefits were altered. The assumption that women were being prescribed bisphosphanates was removed, and the assumed residual benefit of 5 years linear decline following cessation of bisphophonates was set to zero.

The Assessment Group did not believe that any of these sensitivity analyses markedly affected the conclusions and thus these values were left constant in the main analyses.

- Those associated with time of convergence, mapping VAS to EQ-5D and adverse events

Figure 41 Univariate analyses regarding the assumed time of convergence, the trials used in the VAS to EQ-5D mapping and the inclusion of adverse events associated with treatment.



A sensitivity analysis altered the assumption regarding the time at which the VAS scores was assumed independent of intervention, and the time point at which convergence started. In the foundation model it was assumed that VAS scores were identical at 36 months and started converging at 24 months, whereas the sensitivity analysis assumed 24 and 12 months respectively. The effect of mapping VAS to EQ-5D excluding the INVEST trial was analysed, as was assuming a 0.02 QALY loss associated with BKP and PVP.

The change in convergence assumption noticeably reduced the net benefit of all interventions and was maintained in the sensitivity analyses as was the effect of an assumed (and acknowledged to be arbitrary) 0.02 QALY decrement for PVP and BKP. Whilst the mapping without INVEST favoured BKP, PVP and OPLA, it was assumed that this would not change the conclusions and was omitted.

- Those associated with using the EQ-5D directly from the trials where possible rather than the mapping from VAS to EQ-D

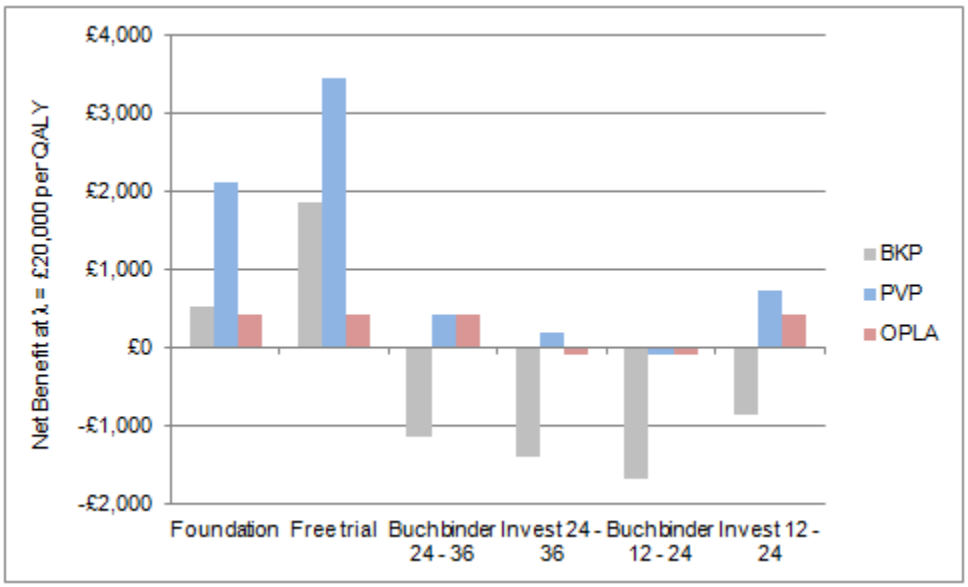
In this sensitivity the difference between the change in EQ-5D in the FREE, Buchbinder and INVEST trials was used directly. Data from Rousing was discarded due to the large baseline difference in EQ-5D. Convergence was tested at both the 24 month and 12 month period for Buchbinder and INVEST. But only at the 24 month period for FREE as the data collection was of 2 years' duration.

For all of the analyses (FREE, Buchbinder and INVEST) it was assumed that the EQ-5D value for BKP would equal the value for PVP. This decision was made given the results from the Liu trial and supported by the midpoint estimate from the network meta-analysis.

When using the FREE data, it was assumed that the values for OPM and OPLA remained at those values mapped from VAS whilst the EQ-5D values for BKP and PVP was estimated as the OPM values plus the difference in change in EQ-5D between BKP and OPM. A limitation of this analysis is that it was assumed that changes to the values for BKP and PVP did not affect the values for OPLA.

For the Buchbinder and INVEST trials the EQ-5D values for OPM and OPLA remained at the VAS mapped values, whilst the EQ-5D value for PVP and BKP was estimated to be the OPLA value plus the difference in change in EQ-5D between PVP and OPLA. A potential limitation of these analyses is that it was assumed that the BKP and PVP values were reduced to nearer the OPLA and OPM network meta-analyses values rather than increasing the OPLA value to the nearer the BKP and PVP meta-analyses values. The methodology reduces the difference in EQ-5D between BKP and PVP compared with OPM and reduces the apparent cost effectiveness of BKP and PVP compared with OPM. The ICERs between BKP / PVP and OPM using the alternative method would be the same as produced in Scenario 5. It is stressed that the ICERs between BKP / PVP and OPLA are independent of the method chosen.

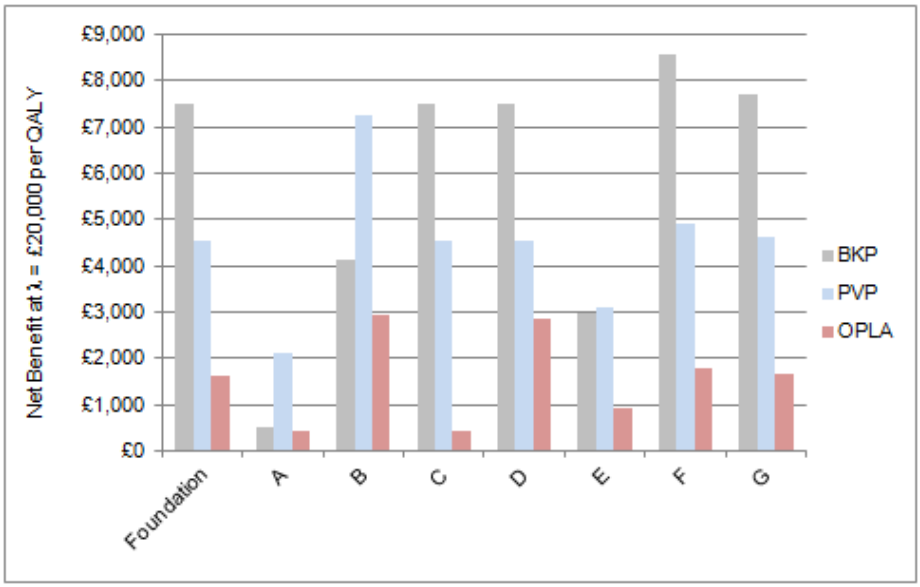
Figure 42 Univariate analyses using change in EQ-5D data directly from trials



All three analyses potentially affected the conclusion, with the net benefit for BKP becoming positive when applying the data from FREE; when the Buchbinder and INVEST trials were used the net benefit difference between PVP and OPLA was noticeably reduced. Given these results the Assessment Group decided to explore the impacts of using the three studies within the full analysis.

- Those affecting mortality or fracture rates

Figure 43 Univariate analyses regarding mortality and fracture rates



Scenario Legend

A: No mortality benefit for any treatment (the foundation analysis in the earlier figures)
B: A pooled mortality benefit for BKP and PVP
C: No mortality benefit for OPLA
D: The mortality benefit for OPLA set equal to PVP
E: No increased mortality risk following the initial vertebral fracture
F: No waning period of the increased mortality risk following the initial fracture
G: No increased risk of mortality in the year of additional vertebral fractures.

This scenario assumed that there was a differential rate in mortality with a hazard ratio of [redacted] for BKP, [redacted] for PVP and [redacted] for OPLA. Other scenarios assumed: no mortality benefit for any treatment; a pooled value of [redacted] for both BKP and PVP ([redacted] for OPLA); hazard ratios of 1 and [redacted] for OPLA; no increased mortality risk following the initial vertebral fracture; no wane time after five years increased mortality risk following the initial vertebral fracture; and no increased risk of fracture in the year of additional vertebral fractures.

It is seen that the assumed mortality effect is a key driver of the results. Removing this for all interventions resulted in PVP having a greater net benefit than BKP; as did assuming that the mortality effects of PVP and BKP were identical. Assuming that the patients did not have a higher risk of mortality for five years due to the prevalent vertebral fracture also resulted in PVP having a higher net benefit than BKP, because the differential mortality benefit of BKP was now applied to a lower underlying rate of mortality. All of the sensitivity analyses performed with the exception of F and G were deemed worthy of additional exploration in the full analyses.

Conclusions from the exploratory univariate analyses.

It is clear that whether or not the interventions have a mortality benefit (and the extent of this if a benefit is assumed) has a considerable effect on the relative cost-effectiveness of the strategies. Additionally the conclusions appeared to be influenced by whether the EQ-5D data were mapped from VAS or taken directly from the trials. These combinations are the six scenarios defined in Figure 36.

Additional sensitivity analyses conducted on these scenarios were: assuming a bed day cost of £0 to set hospitalisation costs equal; altering the assumed cost of equipment for OPLA and the cost of the procedure; altering the time of convergence; and including potential QALY losses associated with adverse events.

Full results for each of the six sensitivity analyses.

The six scenarios are shown diagrammatically in Figure 36. The results from each are discussed in turn. Each scenario is subjected to sensitivity analysis exploring the impacts of changes to the following assumptions: assuming a bed day cost of £0 to set hospitalisation costs equal; altering the assumed cost of equipment for OPLA and the cost of the procedure; altering the time of convergence; and including potential QALY losses associated with adverse events. It is noted that combinations of these sensitivity analyses may represent arguably more plausible central estimates of the cost-effectiveness of the interventions than the unadulterated scenarios and should be provided with equal weight. For example, assuming that the costs of OPLA are identical to the costs of PVP is likely to be favourable to PVP when a comparison with OPLA is made. In analyses where both BKP and PVP have a cost per QALY gained value greater than £20,000, a figure in parenthesis denotes the cost per QALY gained with OPLA removed.

However, for brevity, plots of the cost-effectiveness plane and the cost-effectiveness acceptability curves have been provided only for the unadulterated scenario. When running the probabilistic scenarios it was noted that the model was non-linear. This was due to the assumed distribution for the increased risk of mortality following fracture which was a hazard ratio of 4.40 (1.85 - 10.60)[363]. The mean of this lognormal distribution is 4.86 which increased the risks of dying for all interventions. As such, the sensitivity analyses are presented having undertaken probabilistic sensitivity analyses.

The results for PVP have been estimated assuming the use of low viscosity cement. The Assessment Group's assumed cost of low viscosity cement was £800 per operation, whilst the cost of high-viscosity cement was £1,546, resulting in an estimated increase of £746 per operation associated with the use of high-viscosity cement. Exploratory analyses of assuming high-viscosity cement for all patients have been undertaken.

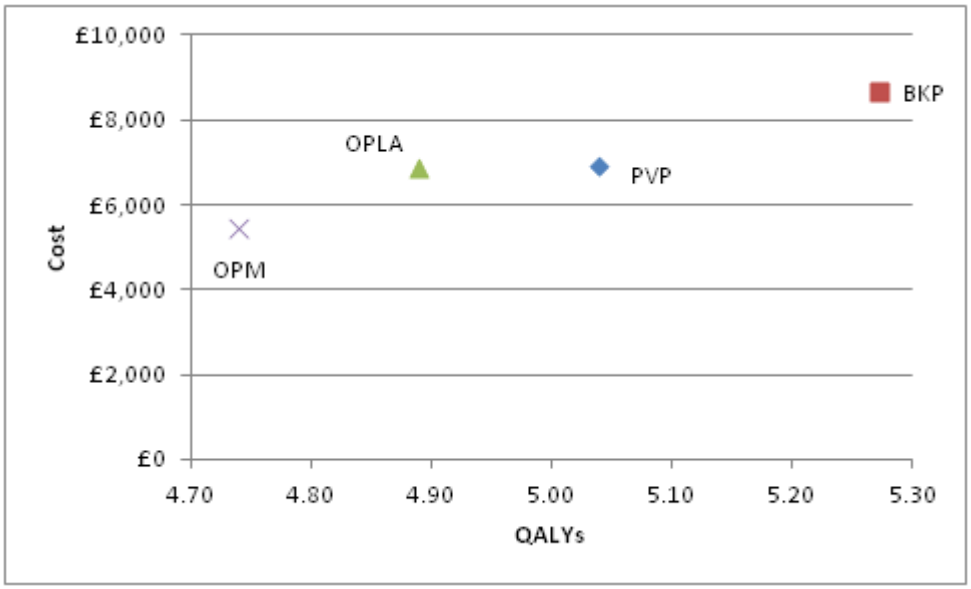
Scenario 1. Differential beneficial effects on mortality assumed for BKP and PVP, utility gain estimated via mapping of stable VAS

The deterministic results are presented in Table 63, with the cost-effectiveness plane depicted in Figure 44

Table 63 The deterministic results produced by Assessment Group - Scenario 1

Intervention	Costs	QALYs	ICER
OPM	£5459	4.74	
OPLA	£6850	4.89	Ext Dominated
PVP	£6897	5.04	£4802
BKP	£8651	5.27	£7488

Figure 44. A plot of the deterministic results produced by Assessment Group - Scenario 1



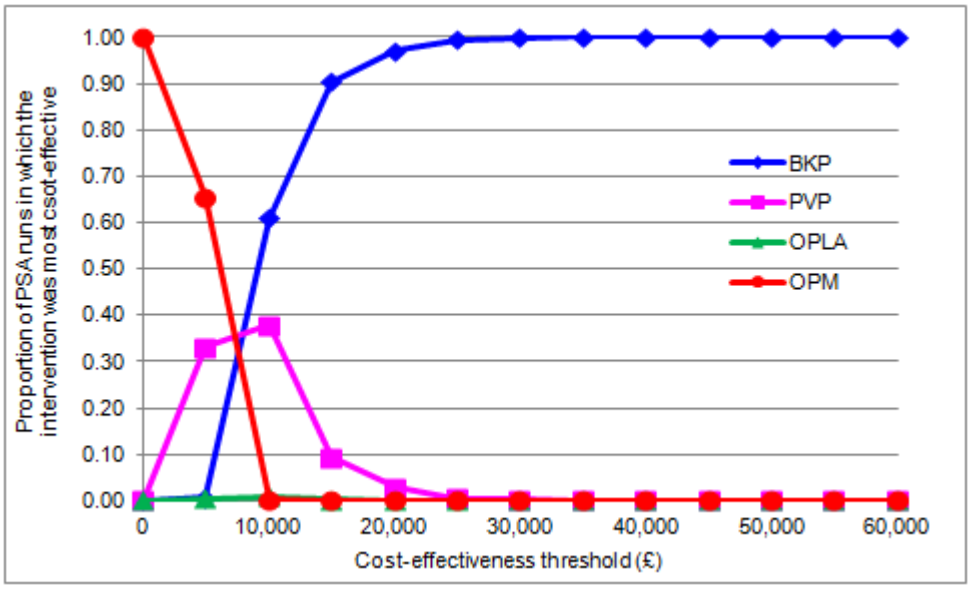
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 64, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 45)

Table 64 The probabilistic results produced by Assessment Group - Scenario 1

Intervention	Costs	QALYs	ICER
OPM	£5399	3.89	
OPLA	£6790	4.03	Ext Dominated
PVP	£6835	4.16	£5341
BKP	£10,147	4.35	£9154

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 45. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 1



Sensitivity analyses conducted on Scenario 1

Table 65 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 65 Sensitivity analyses conducted on Scenario 1

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3916	3.89	
OPLA	£5351	4.03	Ext Dominated
PVP	£5396	4.16	£8184
BKP	£7400	4.35	£10,490
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5399	3.89	
OPLA	£5815	4.03	£3054
PVP	£6835	4.16	£7684
BKP	£8584	4.35	£9154
The probabilistic results produced when it was assumed that convergence of EQ-5D scores began at 12 months and were equal at 24 months			
OPM	£5399	3.95	
OPLA	£6790	4.06	Ext Dominated
PVP	£6835	4.17	£6730
BKP	£8584	4.36	£9142
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5339	3.89	
OPLA	£6790	4.03	Ext Dominated
PVP	£6835	4.14	£5771
BKP	£8584	4.33	£9154
All of the above sensitivity analyses combined			
OPM	£3196	3.95	
OPLA	£4376	4.06	£10,672
PVP	£5396	4.15	Ext Dominated
BKP	£7400	4.34	£11,033

Scenario 2. Differential beneficial effects on mortality assumed for BKP and PVP, utility gain estimated via trials reporting EQ-5D

These analyses have been subdivided into three categories based on whether the FREE data, the Buchbinder et al data or the INVEST data were used.

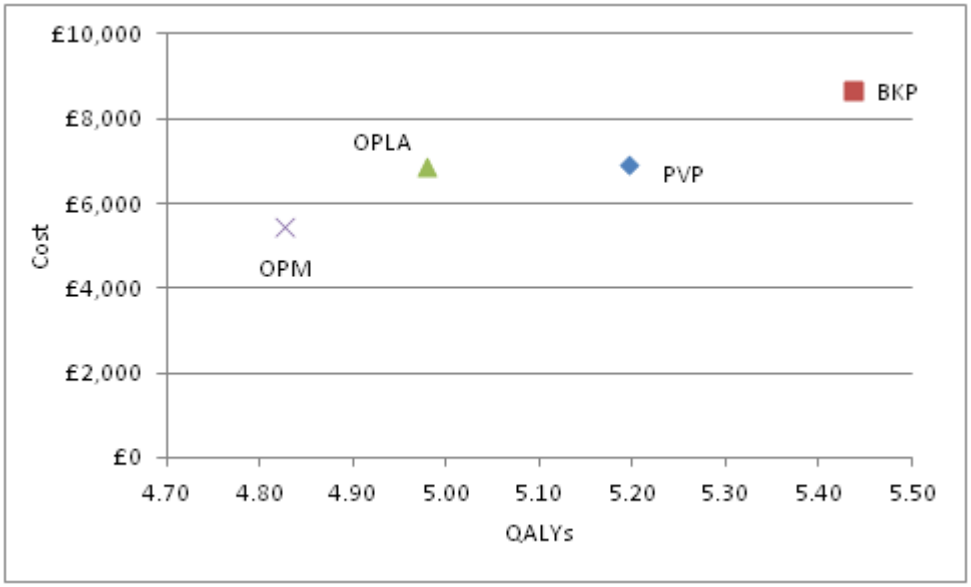
Analyses using the FREE data

The deterministic results are presented in Table 66, with the cost-effectiveness plane depicted in Figure 46

Table 66 The deterministic results produced by Assessment Group - Scenario 2: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5459	4.83	
OPLA	£6850	4.98	Ext Dominated
PVP	£6897	5.20	£3892
BKP	£8651	5.44	£7289

Figure 46. A plot of the deterministic results produced by Assessment Group - Scenario 2: FREE data



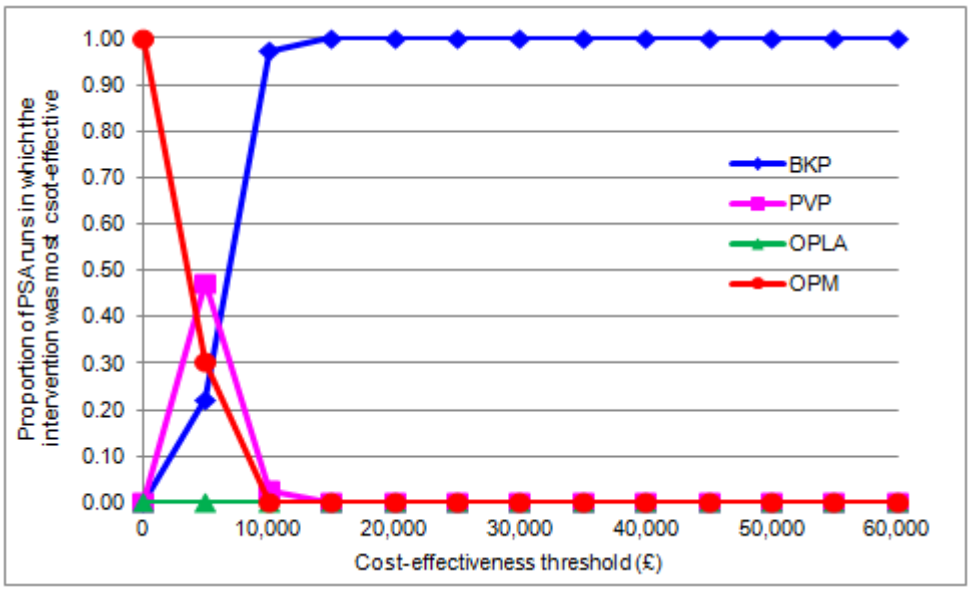
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 67, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 47)

Table 67 The probabilistic results produced by Assessment Group - Scenario 2: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5339	4.75	
OPLA	£6790	4.90	Ext Dominated
PVP	£6835	5.05	£4806
BKP	£8584	5.35	£5814

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 47. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 FREE data



Sensitivity analyses conducted on Scenario 2: FREE Data

Table 68 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 68 Sensitivity analyses conducted on Scenario 2: FREE Data

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3916	4.75	
OPLA	£5351	4.90	Ext Dominated
PVP	£5396	5.05	Ext Dominated
BKP	£7400	5.35	£7012
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5339	4.75	
OPLA	£5815	4.90	£2766
PVP	£6835	5.05	Ext Dominated
BKP	£8584	5.35	£6163
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5339	4.75	
OPLA	£6790	4.90	Ext Dominated
PVP	£6835	5.03	£5151
BKP	£8584	5.33	£5814
All of the above sensitivity analyses combined			
OPM	£3196	4.75	
OPLA	£4376	4.90	Ext Dominated
PVP	£5396	5.03	Ext Dominated
BKP	£7400	5.33	£7254

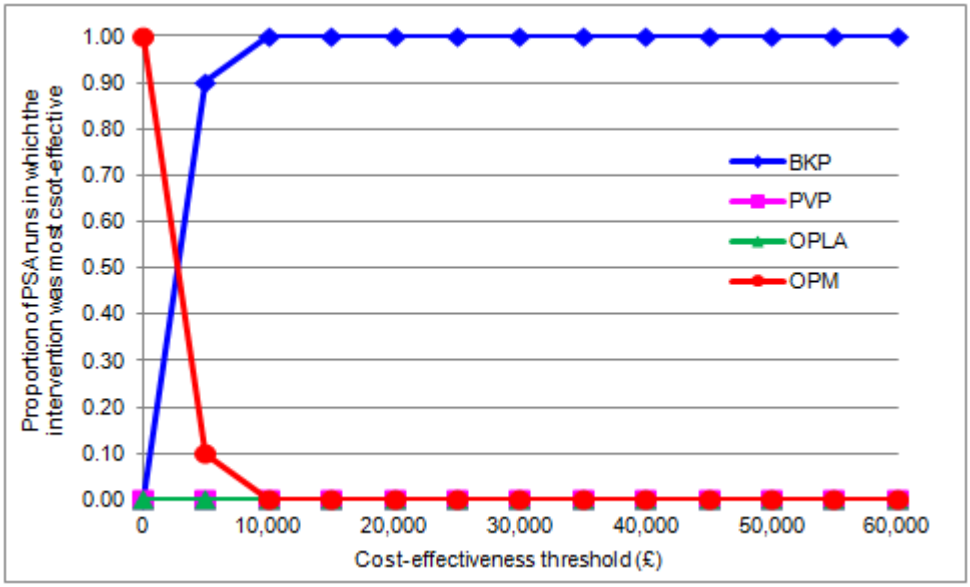
Analyses using the Buchbinder et al data

The deterministic results are presented in Table 69 with the cost-effectiveness plane depicted in Figure 48

Table 69 The deterministic results produced by Assessment Group - Scenario 2: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA	£6850	4.88	Ext Dominated	£6850	4.88	Ext Dominated
PVP	£6897	4.94	Ext Dominated	£6897	4.94	£6703
BKP	£8651	5.17	£7572	£8651	5.17	£7526

Figure 48. A plot of the deterministic results produced by Assessment Group - Scenario 2: Buchbinder data convergence starts at 24 months



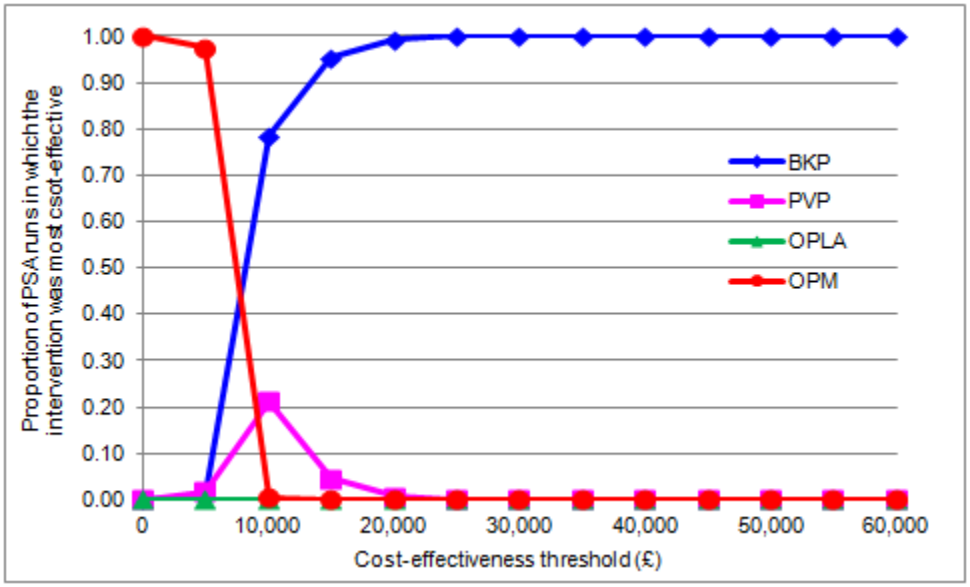
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 70, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 49)

Table 70 The probabilistic results produced by Assessment Group Scenario - 2: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6790	4.79	Ext Dominated	£6790	4.80	Ext Dominated
PVP	£6835	4.86	Ext Dominated	£6835	4.86	£6818
BKP	£8584	5.08	£7724	£8584	5.08	£7688

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 49. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 Buchbinder data



Sensitivity analyses conducted on Scenario 2: Buchbinder Data

Table 71 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 71 Sensitivity analyses conducted on Scenario 2: Buchbinder Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£5351	4.79	Ext Dominated	£5351	4.80	Ext Dominated
PVP	£5396	4.86	Ext Dominated	£5396	4.86	Ext Dominated
BKP	£7400	5.08	£10,196	£7400	5.08	£9597
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5815	4.79	£3378	£5815	4.80	£2794
PVP	£6835	4.86	Ext Dominated	£6835	4.86	Ext Dominated
BKP	£8584	5.08	£9572	£8584	5.08	£9571
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6790	4.79	Ext Dominated	£6790	4.80	Ext Dominated
PVP	£6835	4.84	Ext Dominated	£6835	4.84	£7534
BKP	£8584	5.06	£8117	£8584	5.06	£7688
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4376	4.79	£9590	£4376	4.80	£7932
PVP	£5396	4.84	Ext Dominated	£5396	4.84	Ext Dominated
BKP	£7400	5.06	£11,230	£8963	5.06	£11,229

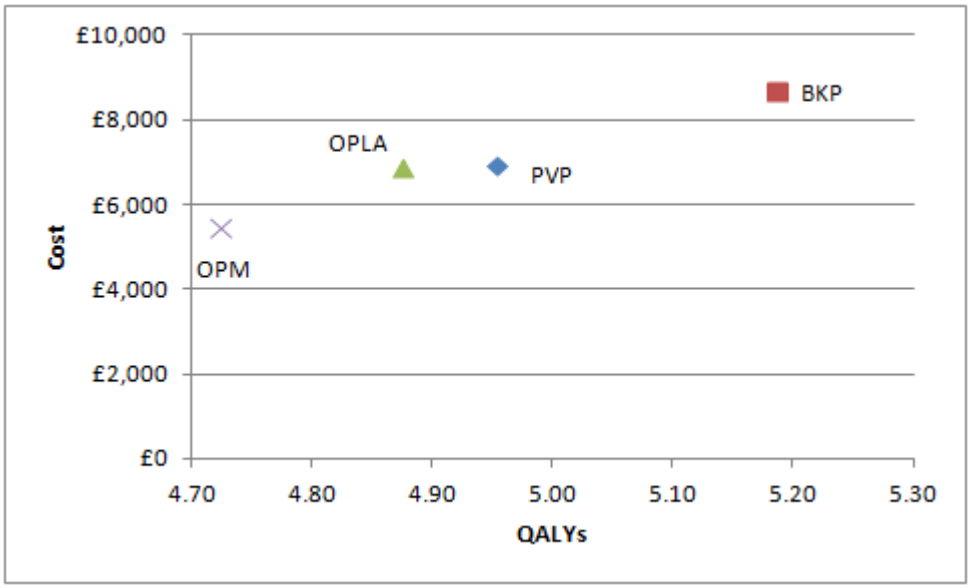
Analyses using the INVEST et al data

The deterministic results are presented in Table 72, with the cost-effectiveness plane depicted in Figure 50

Table 72 The deterministic results produced by Assessment Group - Scenario 2: INVEST data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA	£6850	4.88	Ext Dominated	£6850	4.88	Ext Dominated
PVP	£6897	4.95	£7084	£6897	4.95	£6279
BKP	£8651	5.19	£7525	£8651	5.19	£7534

Figure 50. A plot of the deterministic results produced by Assessment Group - Scenario 2: INVEST data convergence starts at 24 months



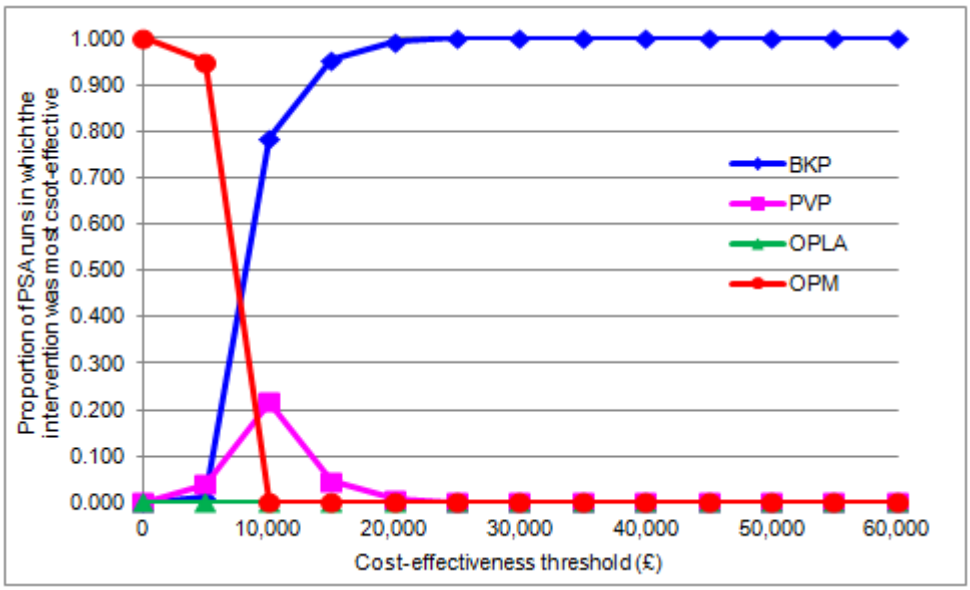
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 73, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 51)

Table 73 The probabilistic results produced by Assessment Group - Scenario 2: INVEST data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6790	4.79	Ext Dominated	£6790	4.80	Ext Dominated
PVP	£6835	4.87	£7205	£6835	4.87	£6381
BKP	£8584	5.10	£7687	£8584	5.10	£7686

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 51. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 2 INVEST data



Sensitivity analyses conducted on Scenario 2: INVEST Data

Table 74 detail the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 74 Sensitivity analyses conducted on Scenario 2: INVEST Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£5351	4.79	Ext Dominated	£5351	4.80	Ext Dominated
PVP	£5396	4.87	Ext Dominated	£5396	4.87	Ext Dominated
BKP	£7400	5.10	£9850	£7400	5.10	£9290
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5815	4.79	£3378	£5815	4.80	£2794
PVP	£6835	4.87	Ext Dominated	£6835	4.87	Ext Dominated
BKP	£8584	5.10	£9115	£8584	5.10	£7037
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6790	4.79	Ext Dominated	£6790	4.80	Ext Dominated
PVP	£6835	4.85	Ext Dominated	£6835	4.85	£7004
BKP	£8584	5.08	£7829	£8584	5.08	£7686
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4376	4.79	£9590	£4376	4.80	£7932
PVP	£5396	4.86	Ext Dominated	£5396	4.84	Ext Dominated
BKP	£7400	5.08	£10,657	£7400	5.06	£10,656

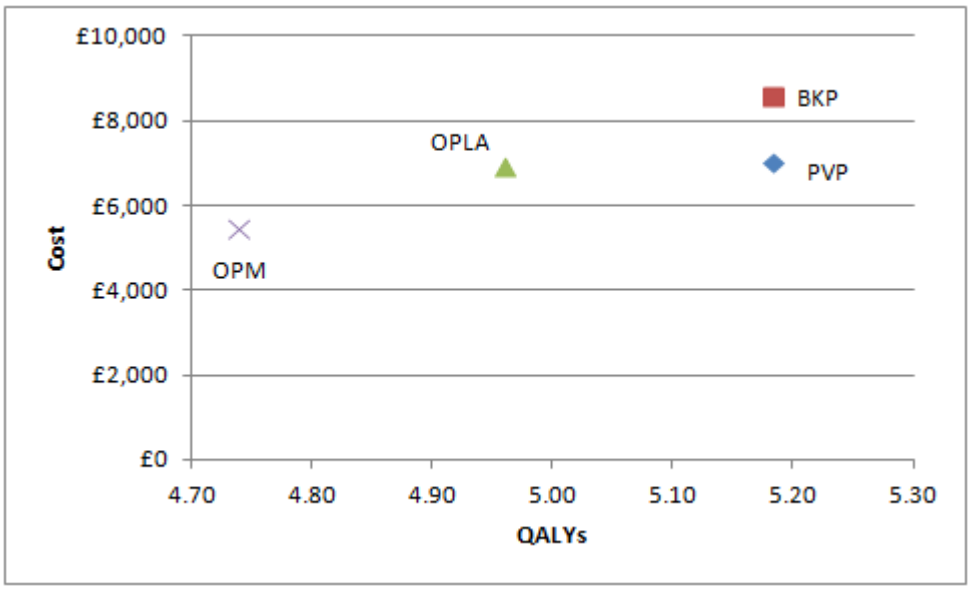
Scenario 3. Equal beneficial effects on mortality assumed for BKP and PVP, utility gain estimated via mapping of stable VAS

The deterministic results are presented in Table 75, with the cost-effectiveness plane depicted in Figure 52

Table 75 The deterministic results produced by Assessment Group - Scenario 3

Intervention	Costs	QALYs	ICER
OPM	£5459	4.74	
OPLA	£6902	4.96	Ext Dominated
PVP	£7002	5.18	£3471
BKP	£8586	5.18	Dominated

Figure 52. A plot of the deterministic results produced by Assessment Group - Scenario 3

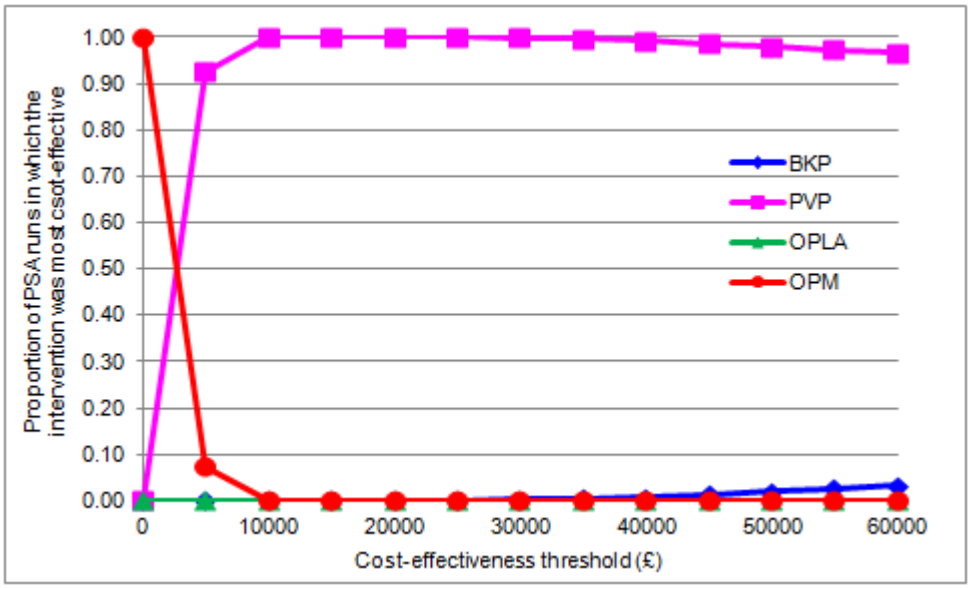


Probabilistic sensitivity analyses were conducted. These results are detailed in Table 76, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 53)

Table 76 The probabilistic results produced by Assessment Group - Scenario 3

Intervention	Costs	QALYs	ICER
OPM	£5399	3.89	
OPLA	£6840	4.09	Ext Dominated
PVP	£6937	4.28	£3969
BKP	£8521	4.28	Dominated

Figure 53. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 3



Sensitivity analyses conducted on Scenario 3

Table 77 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 77 Sensitivity analyses conducted on Scenario 3

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3196	3.89	
OPLA	£5401	4.09	Ext Dominated
PVP	£5498	4.28	£5941
BKP	£7337	4.28	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5399	3.89	
OPLA	£5865	4.09	£2405
PVP	£6937	4.28	£5529
BKP	£8521	4.28	Dominated
The probabilistic results produced when it was assumed that convergence of EQ-5D scores began at 12 months and were equal at 24 months			
OPM	£5399	3.95	
OPLA	£6840	4.12	Ext Dominated
PVP	£6937	4.28	£4,632
BKP	£8521	4.28	Dominated
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5399	3.89	
OPLA	£6840	4.09	Ext Dominated
PVP	£6937	4.26	£4185
BKP	£8521	4.26	Dominated
All of the above sensitivity analyses combined			
OPM	£3196	3.95	
OPLA	£4425	4.12	£7308
PVP	£5498	4.26	£7458
BKP	£7337	4.26	Dominated
As above plus mortality effect of OPLA set to equal BKP and PVP			
OPM	£3196	3.95	
OPLA	£4523	4.23	£4,723
PVP	£5498	4.26	£31,304 (£7377)
BKP	£7337	4.26	Dominated (Dominated)

Numbers in parentheses indicate the ICER if OPLA were not considered to be a comparator

Scenario 4. Equal beneficial effects assumed for BKP and PVP, utility gain estimated via trials reporting EQ-5D

These analyses have been subdivided into three categories based on whether the FREE data, the Buchbinder et al data or the INVEST data were used.

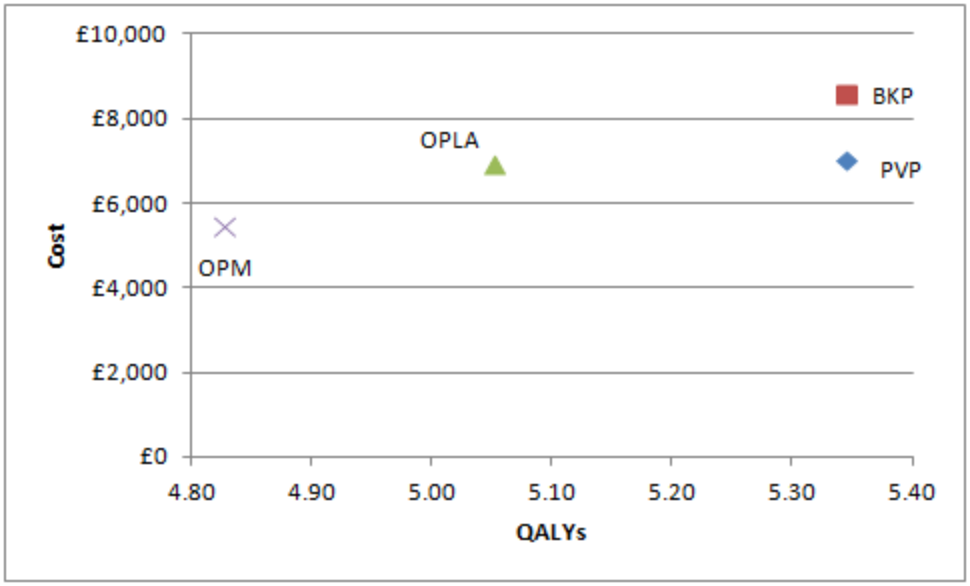
Analyses using the FREE data

The deterministic results are presented in Table 78, with the cost-effectiveness plane depicted in Figure 54

Table 78 The deterministic results produced by Assessment Group - Scenario 4: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5459	4.83	
OPLA	£6902	5.05	Ext Dominated
PVP	£7002	5.35	£2977
BKP	£8586	5.35	Dominated

Figure 54. A plot of the deterministic results produced by Assessment Group - Scenario 4: FREE data



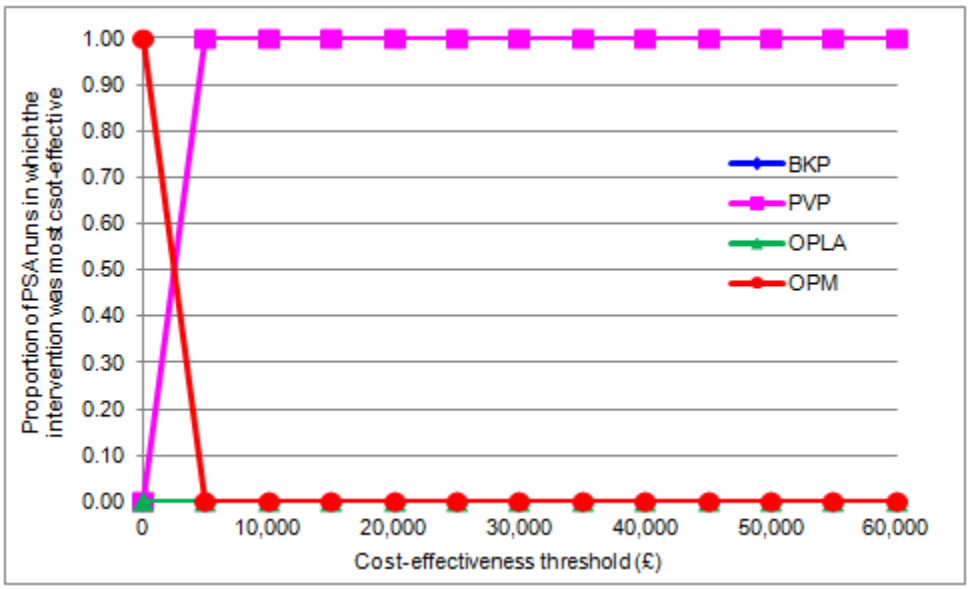
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 76, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 55)

Table 79 The probabilistic results produced by Assessment Group - Scenario 4: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5339	4.75	
OPLA	£6840	4.97	Ext Dominated
PVP	£6937	5.26	£3015
BKP	£8521	5.26	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 55. The cost-effectiveness acceptability curve produced by Assessment Group -Scenario 4 FREE data



Sensitivity analyses conducted on Scenario 4: FREE Data

Table 80 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 80 Sensitivity analyses conducted on Scenario 4: FREE Data

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3196	4.75	
OPLA	£5401	4.97	Ext Dominated
PVP	£5498	5.26	£4513
BKP	£7337	5.26	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5339	4.75	
OPLA	£5865	4.97	£2109
PVP	£6937	5.26	£3705
BKP	£8521	5.26	Dominated
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5339	4.75	
OPLA	£6840	4.97	Ext Dominated
PVP	£6937	5.24	£3138
BKP	£8521	5.24	Dominated
All of the above sensitivity analyses combined			
OPM	£3196	4.75	
OPLA	£4425	4.97	Ext Dominated
PVP	£5498	5.24	£4697
BKP	£7337	5.24	Dominated
All of the above sensitivity analyses combined plus mortality effect of OPLA set to equal BKP and PVP			
OPM	£3196	4.75	
OPLA	£4523	5.10	£3705
PVP	£5498	5.24	£7386
BKP	£7337	5.24	Dominated

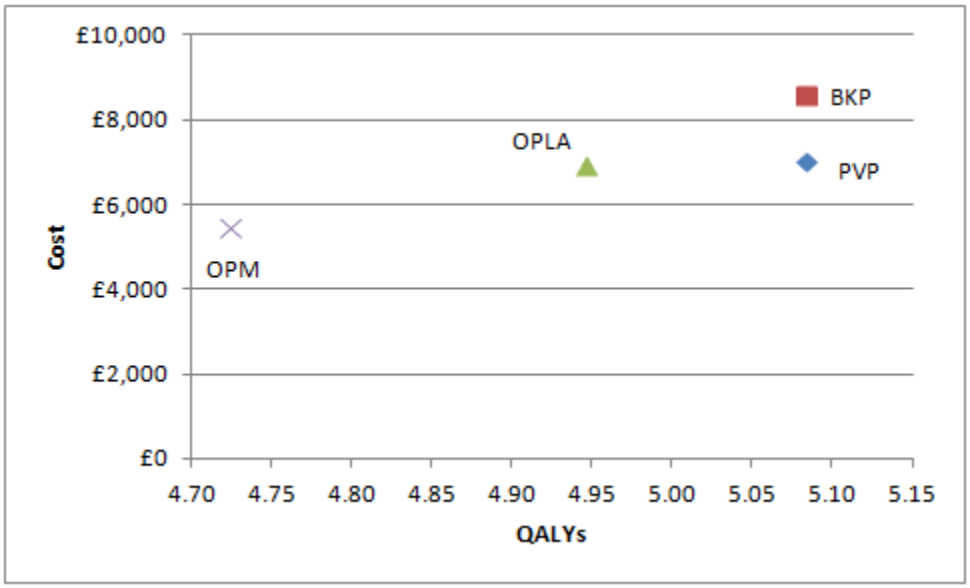
Analyses using the Buchbinder et al data

The deterministic results are presented in Table 81, with the cost-effectiveness plane depicted in Figure 56

Table 81 The deterministic results produced by Assessment Group - Scenario 4: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA	£6902	4.95	Ext Dominated	£6902	4.95	Ext Dominated
PVP	£7002	5.08	£4637	£7002	5.08	£4301
BKP	£8586	5.08	Dominated	£8586	5.08	Dominated

Figure 56. A plot of the deterministic results produced by Assessment Group - Scenario 4: Buchbinder data convergence starts at 24 months



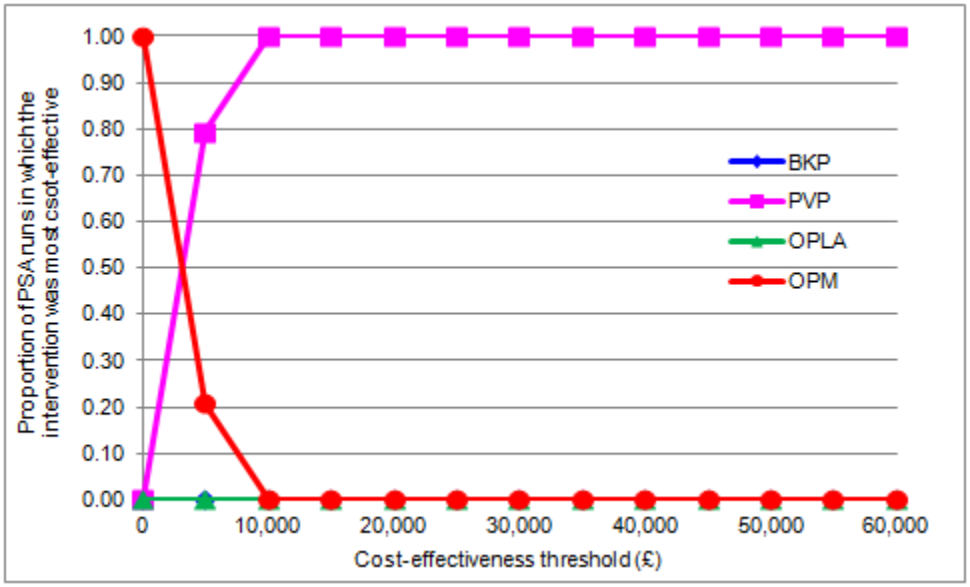
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 82, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 57)

Table 82 The probabilistic results produced by Assessment Group - Scenario 4: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6840	4.86	Ext Dominated	£6840	4.86	Ext Dominated
PVP	£6937	5.00	£4720	£6937	5.00	£4375
BKP	£8521	5.00	Dominated	£8521	5.00	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 57. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 4 Buchbinder data



Sensitivity analyses conducted on Scenario 4: Buchbinder Data

Table 83 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 83 Sensitivity analyses conducted on Scenario 4: Buchbinder Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£5401	4.86	Ext Dominated	£5401	4.86	Ext Dominated
PVP	£5498	5.00	£7065	£5498	5.00	£6548
BKP	£7337	5.00	Dominated	£7337	5.00	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5865	4.86	£2427	£5865	4.86	£2140
PVP	£6937	5.00	£7997	£6937	5.00	£7997
BKP	£8521	5.00	Dominated	£8521	5.00	Dominated
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6840	4.86	Ext Dominated	£6840	4.86	Ext Dominated
PVP	£6937	4.98	£5029	£6937	4.98	£4639
BKP	£8521	4.98	Dominated	£8521	4.98	Dominated
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4425	4.86	£6413	£4425	4.86	£5654
PVP	£5498	4.98	£9399	£5498	4.98	£9398
BKP	£7337	4.98	Dominated	£7337	4.98	Dominated

Table 83 (continued) Sensitivity analyses conducted on Scenario 4: Buchbinder Data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
All of the above sensitivity analyses combined plus mortality effect of OPLA set to equal BKP and PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4523	5.00	£4071	£4523	5.00	£3773
PVP	£5498	5.00	Dominated (£7527)	£5498	5.00	Dominated (£6943)
BKP	£7337	5.00	Dominated (Dominated)	£7337	5.00	Dominated (Dominated)

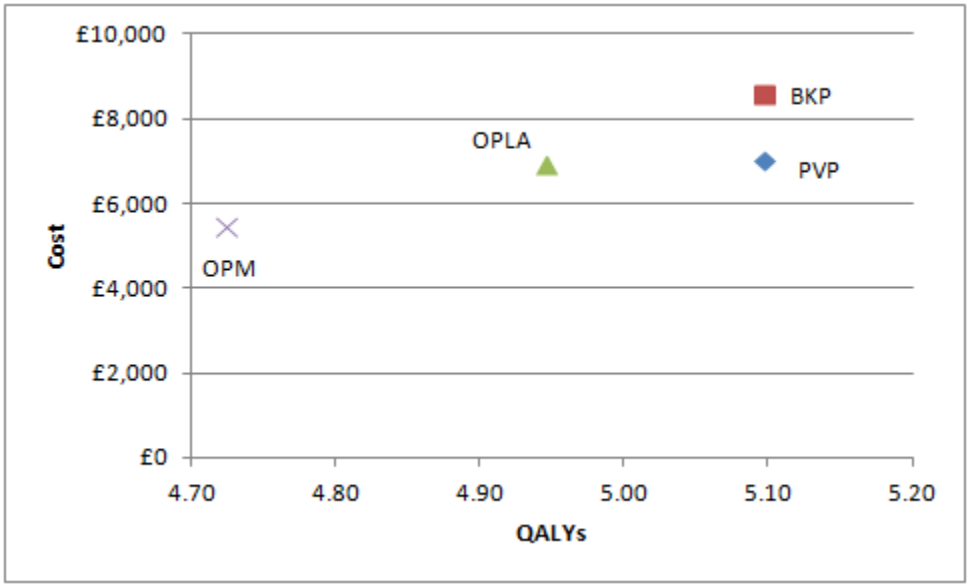
Analyses using the INVEST et al data

The deterministic results are presented in Table 84, with the cost-effectiveness plane depicted in Figure 58

Table 84 The deterministic results produced by Assessment Group - Scenario 4: INVEST data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA	£6902	4.95	Ext Dominated	£6902	4.95	Ext Dominated
PVP	£7002	5.10	£4444	£7002	5.10	£4134
BKP	£8586	5.10	Dominated	£8586	5.10	Dominated

Figure 58. A plot of the deterministic results produced by Assessment Group - Scenario 4: INVEST data convergence starts at 24 months



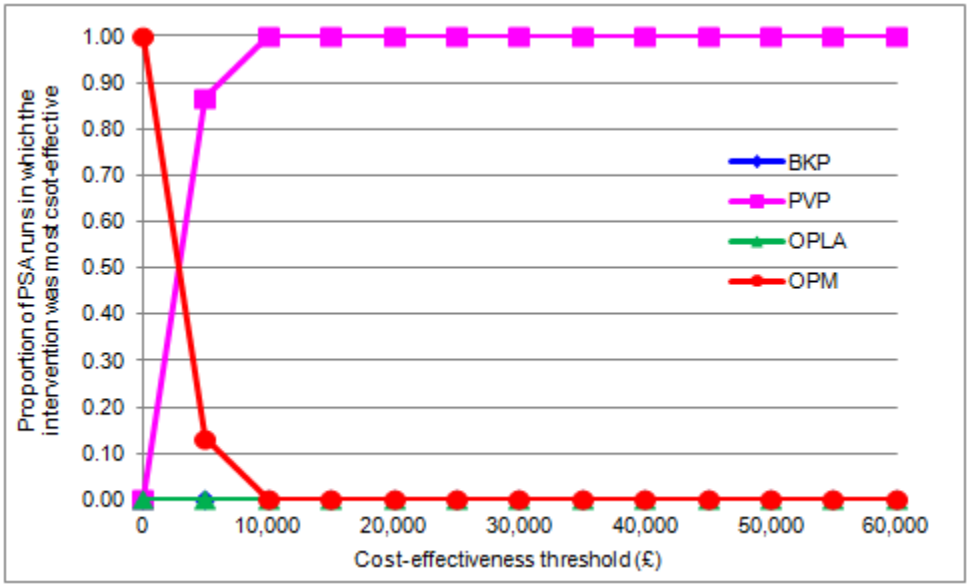
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 85, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 59)

Table 85 The probabilistic results produced by Assessment Group Scenario 4: INVEST data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6840	4.86	Ext Dominated	£6840	4.86	Ext Dominated
PVP	£6937	5.01	£4520	£6937	5.01	£4202
BKP	£8521	5.01	Dominated	£8521	5.01	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 59. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 4 INVEST data



Sensitivity analyses conducted on Scenario 4: INVEST Data

Table 86 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 86 Sensitivity analyses conducted on Scenario 4: INVEST Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£5401	4.86	Ext Dominated	£5401	4.86	Ext Dominated
PVP	£5498	5.01	£6765	£5498	5.01	£6289
BKP	£7337	5.01	Dominated	£7337	5.01	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5865	4.86	£2427	£5865	4.86	£2140
PVP	£6937	5.01	£7219	£6937	5.01	£7219
BKP	£8521	5.01	Dominated	£8521	5.01	Dominated
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6840	4.86	Ext Dominated	£6840	4.86	Ext Dominated
PVP	£6937	4.99	£4802	£6937	4.99	£4445
BKP	£8521	4.99	Dominated	£8521	4.99	Dominated
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.67	
OPLA	£4425	4.86	£6413	£4425	4.86	£5654
PVP	£5498	4.99	£8342	£5498	4.99	£8341
BKP	£7337	4.99	Dominated	£7337	4.99	Dominated

Table 86 (continued) Sensitivity analyses conducted on Scenario 4: INVEST Data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
All of the previous sensitivity analyses combined plus mortality effect of OPLA set to equal BKP and PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4523	5.00	£4071	£4523	5.00	£3773
PVP	£5498	5.00	Dominated (£7187)	£5498	5.00	Dominated (£6652)
BKP	£7337	5.00	Dominated (Dominated)	£7337	5.00	Dominated (Dominated)

Numbers in parentheses indicate the ICER if OPLA were not considered to be a comparator

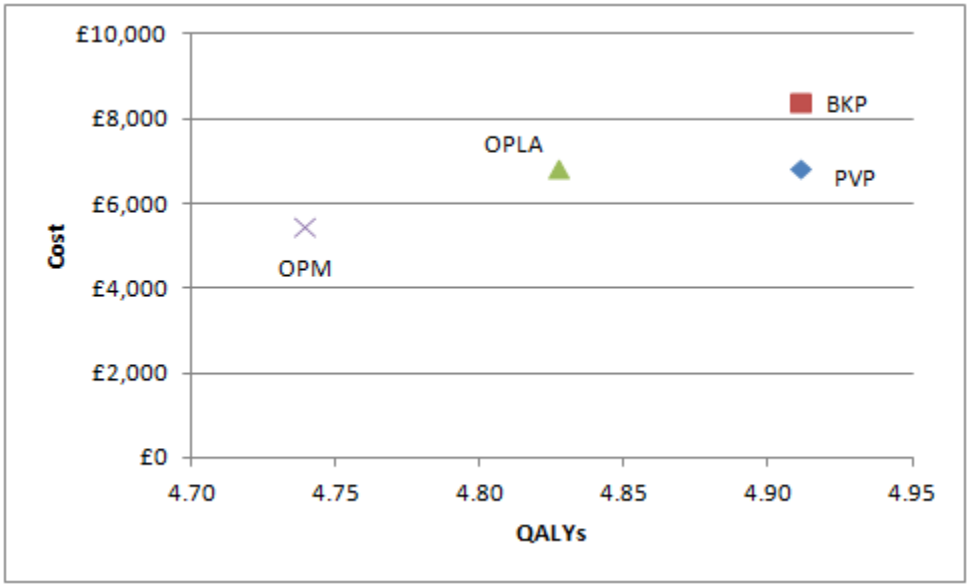
Scenario 5. Differential beneficial effects assumed for BKP and PVP, utility gain estimated via mapping of stable VAS

The deterministic results are presented in Table 87, with the cost-effectiveness plane depicted in Figure 60

Table 87 The deterministic results produced by Assessment Group - Scenario 5

Intervention	Costs	QALYs	ICER
OPM	£5459	4.74	
OPLA	£6804	4.83	Dominated
PVP	£6804	4.91	£7802
BKP	£8388	4.91	Dominated

Figure 60. A plot of the deterministic results produced by Assessment Group - Scenario 5

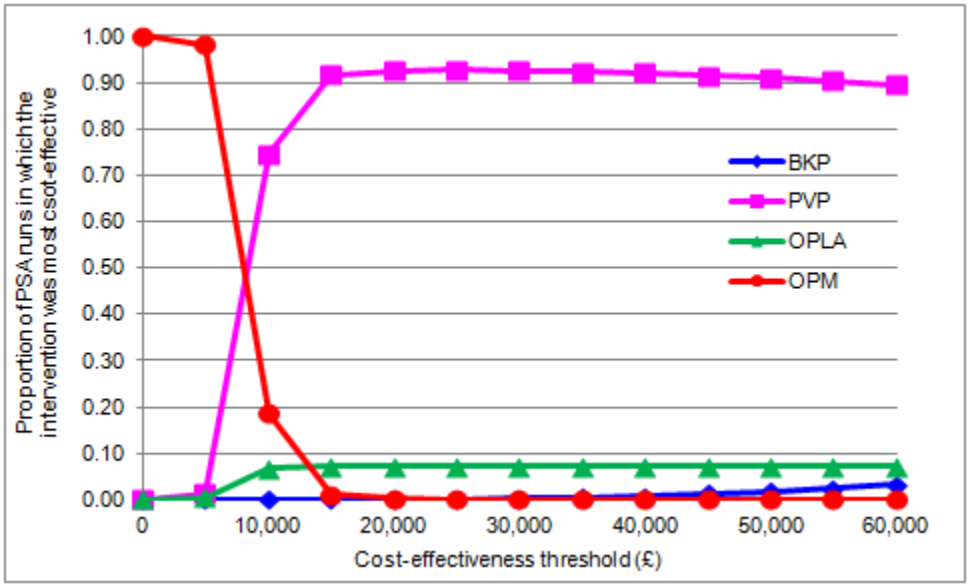


Probabilistic sensitivity analyses were conducted. These results are detailed in Table 88, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 61)

Table 88 The probabilistic results produced by Assessment Group - Scenario 5

Intervention	Costs	QALYs	ICER
OPM	£5399	3.89	
OPLA	£6746	3.98	Dominated
PVP	£6746	4.06	£8139
BKP	£8330	4.06	Dominated

Figure 61. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 5



Sensitivity analyses conducted on Scenario 5:

Table 89 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 89 Sensitivity analyses conducted on Scenario 5

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3196	3.89	
OPLA	£5307	3.98	Dominated
PVP	£5307	4.06	£12,757
BKP	£7146	4.06	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5399	3.89	
OPLA	£5771	3.98	£4362
PVP	£6746	4.06	£12,144
BKP	£8330	4.06	Dominated
The probabilistic results produced when it was assumed that convergence of EQ-5D scores began at 12 months and were equal at 24 months			
OPM	£5399	3.96	
OPLA	£6746	4.01	Dominated
PVP	£6746	4.06	£12,250
BKP	£8330	4.06	Dominated
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5399	3.89	
OPLA	£6746	3.98	Dominated
PVP	£6746	4.04	£9258
BKP	£8330	4.04	Dominated
All of the above sensitivity analyses combined			
OPM	£3196	3.95	
OPLA	£4332	4.01	£19,109
PVP	£5307	4.04	£31,953 (£23,469)
BKP	£7146	4.04	Dominated (Dominated)

Numbers in parenthesis indicate the ICER were OPLA not considered a comparator

Scenario 6. Equal beneficial effects assumed for BKP and PVP, utility gain estimated via trials reporting EQ-5D

These analyses have been subdivided into three categories based on whether the FREE data, the Buchbinder et al data or the INVEST data were used.

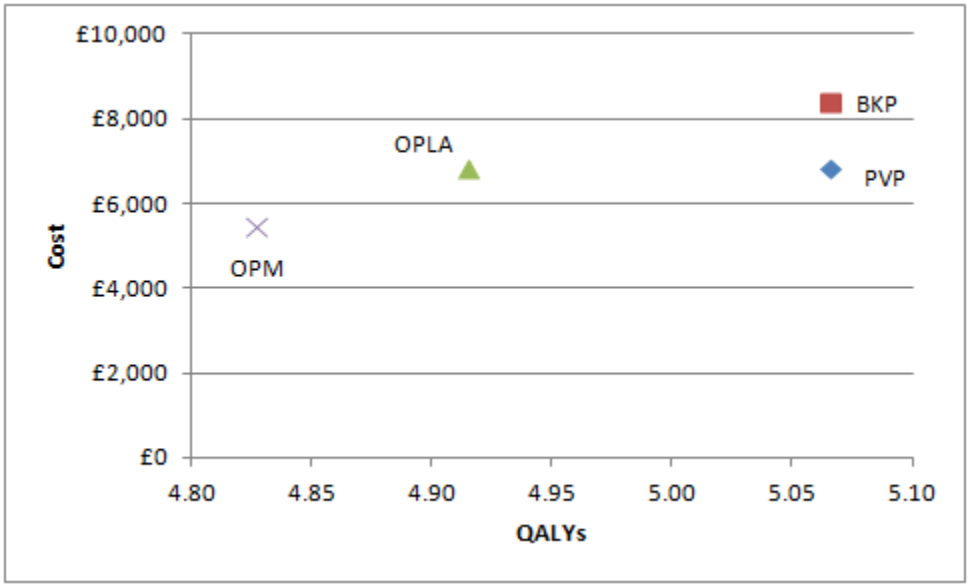
Analyses using the FREE data

The deterministic results are presented in Table 90, with the cost-effectiveness plane depicted in Figure 62

Table 90 The deterministic results produced by Assessment Group Scenario 6: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5459	4.83	
OPLA	£6804	4.92	Ext Dominated
PVP	£6804	5.07	£5625
BKP	£8388	5.07	Dominated

Figure 62. A plot of the deterministic results produced by Assessment Group Scenario 6: FREE data



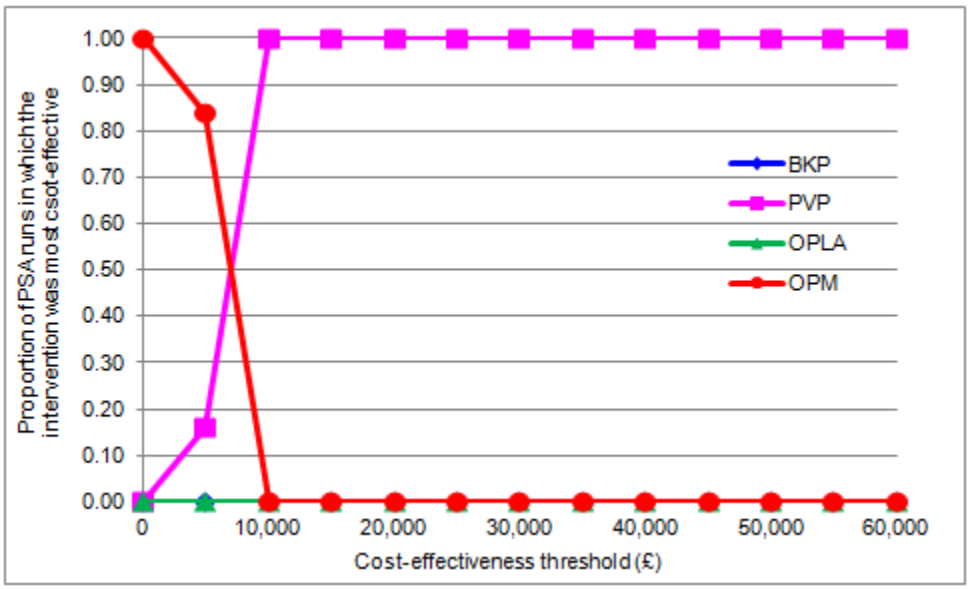
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 91, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve (Figure 63)

Table 91 The probabilistic results produced by Assessment Group - Scenario 6: FREE data

Intervention	Costs	QALYs	ICER
OPM	£5399	4.75	
OPLA	£6746	4.83	Dominated
PVP	£6746	4.98	£5669
BKP	£8330	4.98	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 63. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 FREE data



Sensitivity analyses conducted on Scenario 6: FREE Data

Table 92 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 92 Sensitivity analyses conducted on Scenario 6: FREE Data

Intervention	Costs	QALYs	ICER
The probabilistic results produced when hospitalisation costs were set to £0 per day			
OPM	£3196	4.75	
OPLA	£5307	4.83	Dominated
PVP	£5307	4.98	£8885
BKP	£7146	4.98	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP			
OPM	£5399	4.75	
OPLA	£5711	4.83	£4227
PVP	£6746	4.98	£6514
BKP	£8388	4.98	Dominated
The probabilistic results produced when it was assumed that BKP and PVP were associated with 0.02 QALY loss			
OPM	£5399	4.75	
OPLA	£5711	4.83	Dominated
PVP	£6746	4.96	£6190
BKP	£8388	4.96	Dominated
All of the above sensitivity analyses combined			
OPM	£3196	4.75	
OPLA	£4332	4.83	Dominated
PVP	£5307	4.98	£9701
BKP	£7146	4.98	Dominated

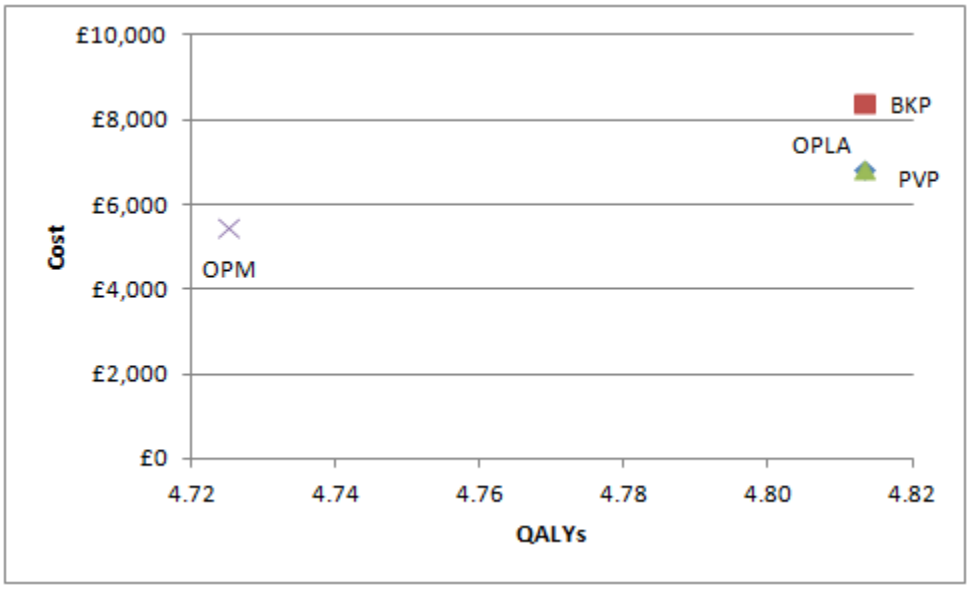
Analyses using the Buchbinder et al data

The deterministic results are presented in Table 93, with the cost-effectiveness plane depicted in Figure 64

Table 93 The deterministic results produced by Assessment Group - Scenario 6: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA / PVP	£6804	4.81	£21,565	£6804	4.81	£15,221
BKP	£8388	4.81	Dominated	£8388	4.81	Dominated

Figure 64. A plot of the deterministic results produced by Assessment Group - Scenario 6: Buchbinder data convergence starts at 24 months



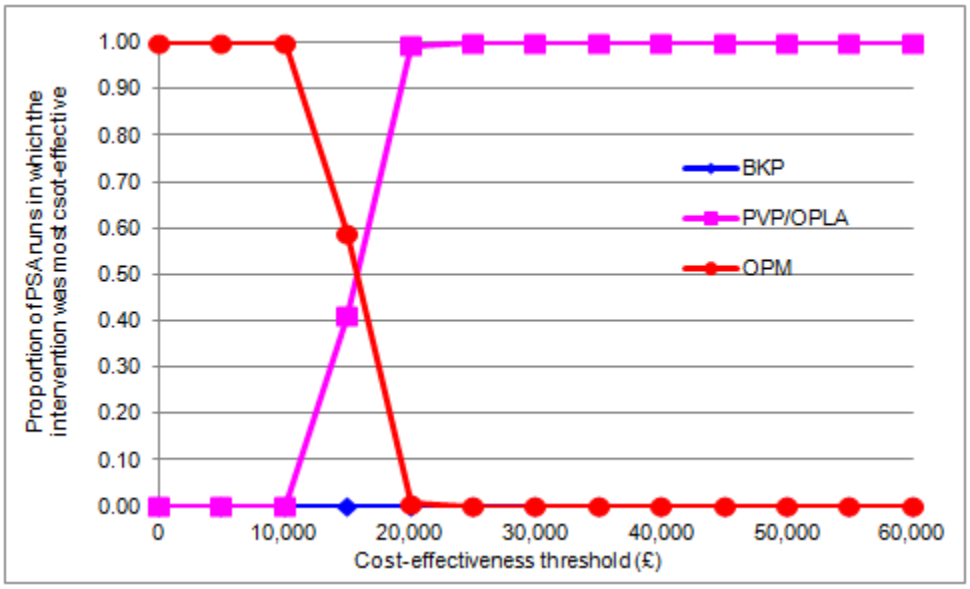
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 94, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 65)

Table 94 The probabilistic results produced by Assessment Group Scenario 6: Buchbinder data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA / PVP	£6746	4.73	£21,670	£6746	4.73	£15,330
BKP	£8330	4.73	Dominated	£8330	4.73	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 65. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 Buchbinder data



Sensitivity analyses conducted on Scenario 6: Buchbinder Data

Table 95 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 95 Sensitivity analyses conducted on Scenario 6: Buchbinder Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA / PVP	£5307	4.73	£33,963	£5401	4.73	£24,027
BKP	£7146	4.73	Dominated	£7146	4.73	Dominated
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5771	4.73	£5975	£5771	4.73	£4227
PVP	£6746	4.73	Dominated (£21,670)	£6746	4.73	Dominated (£15,330)
BKP	£8330	4.73	Dominated (Dominated)	£8330	4.73	Dominated (Dominated)
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6746	4.73	£21,670	£6746	4.73	£15,330
PVP	£6746	4.71	Dominated (£31,950)	£6746	4.71	Dominated (£19,849)
BKP	£8330	4.71	Dominated (Dominated)	£8330	4.71	Dominated (Dominated)

Table 95 (continued) Sensitivity analyses conducted on Scenario 6: Buchbinder Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4332	4.73	£18268	£4332	4.73	£12,924
PVP	£5307	4.71	Dominated (£50,076)	£5307	4.71	Dominated (£31,109)
BKP	£7146	4.71	Dominated (Dominated)	£7146	4.71	Dominated (Dominated)

Numbers in parenthesis indicate the ICER if OPLA were not considered a comparator

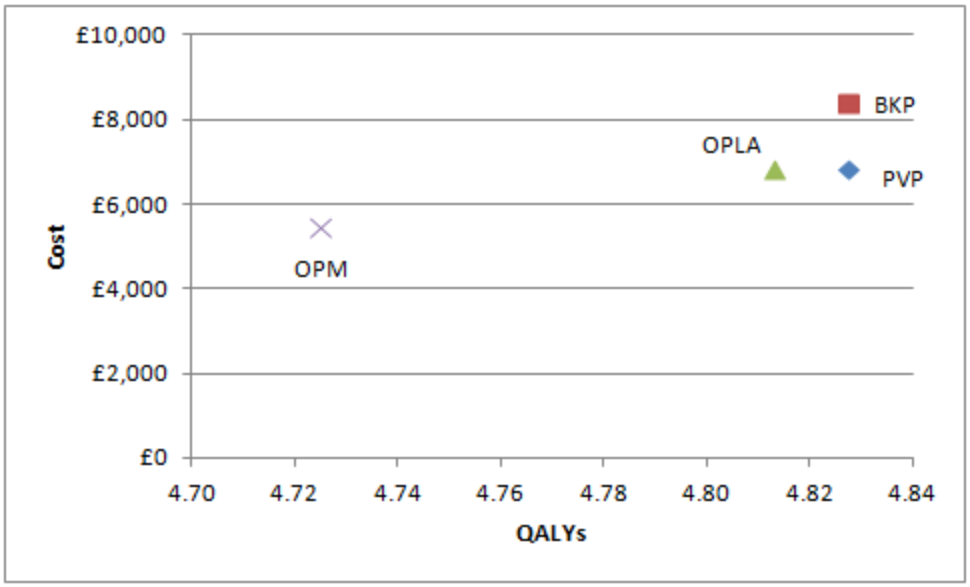
Analyses using the INVEST et al data

The deterministic results are presented in Table 96, with the cost-effectiveness plane depicted in Figure 66

Table 96 The deterministic results produced by Assessment Group - Scenario 6: INVEST data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5459	4.75		£5459	4.73	
OPLA	£6804	4.81	Dominated	£6804	4.81	Dominated
PVP	£6804	4.83	£17,515	£6804	4.83	£13,085
BKP	£8388	4.83	Dominated	£8388	4.83	Dominated

Figure 66. A plot of the deterministic results produced by Assessment Group - Scenario 6: INVEST data convergence starts at 24 months



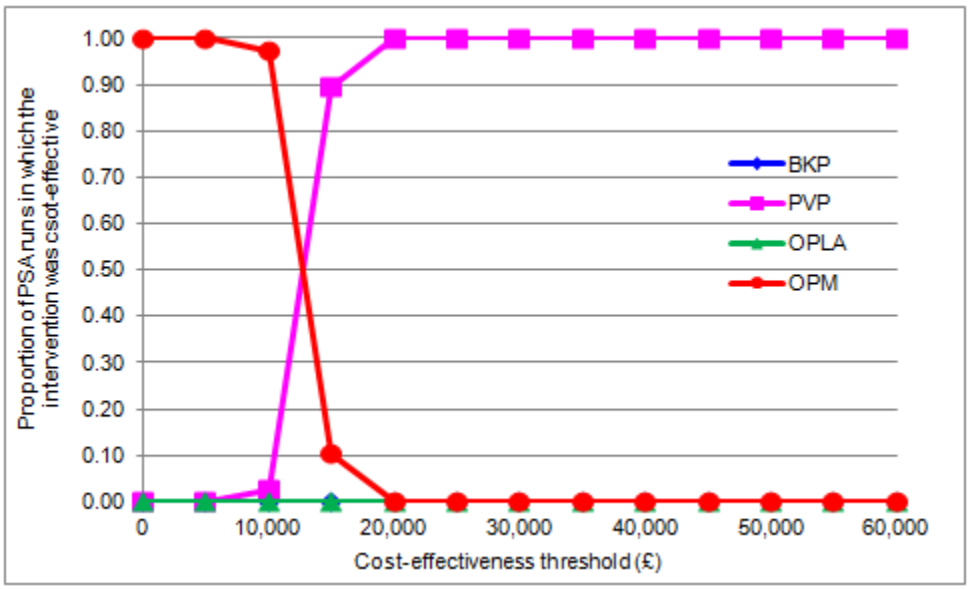
Probabilistic sensitivity analyses were conducted. These results are detailed in Table 97, with an assessment of the uncertainty of the adoption decision displayed in a cost-effectiveness acceptability curve assuming convergence starts at 24 months (Figure 67)

Table 97 The probabilistic results produced by Assessment Group - Scenario 6: INVEST data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6746	4.73	Dominated	£6746	4.73	Dominated
PVP	£6746	4.75	£17,596	£6746	4.75	£13,173
BKP	£8330	4.75	Dominated	£8330	4.75	Dominated

It is seen that the results from the probabilistic sensitivity analyses differed slightly from the deterministic values.

Figure 67. The cost-effectiveness acceptability curve produced by Assessment Group - Scenario 6 INVEST data



Sensitivity analyses conducted on Scenario 6: INVEST Data

Table 98 details the sensitivity analyses conducted by the Assessment Group. Given the non-linearity of the model the results from the probabilistic sensitivity analyses are presented.

Table 98 Sensitivity analyses conducted on Scenario 6: INVEST Data

	Convergence between 12 and 24 months			**Convergence between 24 and 36 months**		
The probabilistic results produced when hospitalisation costs were set to £0 per day						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£5307	4.73	Dominated	£5307	4.73	Dominated
PVP	£5307	4.75	£27,577	£5307	4.75	£20,645
BKP	£7146	4.75	Dominated (Dominated)	£7146	4.75	Dominated (Dominated)
The probabilistic results produced when the cost of the OPLA procedure was set to 50% that of PVP and the cost of OPLA equipment was set to 60% that of PVP						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£5771	4.73	£5975	£5711	4.73	£4227
PVP	£6746	4.75	£67,780 (£17,596)	£6746	4.75	£67,780 (£13,173)
BKP	£8330	4.75	Dominated (Dominated)	£8330	4.75	Dominated (Dominated)
The probabilistic results when it was assumed that BKP and PVP were associated with 0.02 QALY loss						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£5399	4.67		£5399	4.65	
OPLA	£6746	4.73	£21,670	£6746	4.73	£15,330
PVP	£6746	4.73	Dominated (£23,819)	£6746	4.73	Dominated (£16,376)
BKP	£8330	4.73	Dominated (Dominated)	£8330	4.73	Dominated (Dominated)

Table 98 (continued) Sensitivity analyses conducted on Scenario 6: INVEST Data

	Convergence between 12 and 24 months			Convergence between 24 and 36 months		
All of the above sensitivity analyses combined						
Intervention	Costs	QALYs	ICER	Costs	QALYs	ICER
OPM	£3196	4.67		£3196	4.65	
OPLA	£4332	4.73	£18,268	£4332	4.73	£12,924
PVP	£5307	4.73	Dominated (£25,665)	£5307	4.73	Dominated (£37,331)
BKP	£7146	4.73	Dominated (Dominated)	£7146	4.73	Dominated (Dominated)

Interpretation of the results

If differential mortality effects with BKP being more effective than PVP (Scenarios 1 and 2) were assumed then BKP always provided the most QALYs and always below a cost of £20,000 per QALY gained. This was maintained even if the cost of BKP was increased assuming a separate kit was needed for each level (data not shown)

Contrastingly, if it was assumed that the mortality effects of PVP and BKP were identical, with OPLA providing half the benefit, (Scenarios 3 and 4) BKP was estimated to be dominated by PVP providing effectively the same QALYs at a higher cost whilst the cost per QALY gained, with PVP having an ICER of below £10,000 per QALY gained compared to OPM. In the analyses where OPLA was assumed to have an identical mortality benefit to BKP and PVP, a reduced cost than PVP, QALY losses due to adverse events for PVP and the EQ-5D data from the RCTs were used, PVP became dominated by OPLA, although if this was not seen to be an appropriate comparator the ICER of PVP compared with OPM was still below £10,000 per QALY gained.

In the analysis where it was assumed that no intervention provided a mortality benefit compared with OPM (Scenarios 5 and 6) then the conclusions altered dependent on the assumptions made. Without directly using EQ-5D data from the blinded RCTs (Scenario 5) PVP typically provided the most QALYs at a cost per QALY gained below £20,000, with one exception which used assumptions unfavourable to PVP (equal hospitalisation stay costs, a reduced cost of OPLA, QALY losses due to adverse events for PVP and an earlier convergence of EQ-5D scores).

Where the blinded RCT data (Bookbinder and INVEST) (Scenario 6) was used BKP was estimated to be dominated by PVP. PVP was often dominated by OPLA or had an ICER greater than £20,000 per QALY; when OPLA was not considered an appropriate comparator on occasions PVP had an ICER greater than £20,000 compared with OPM. However, it is commented that the methodology of reducing the BKP and PVP estimates to nearer the OPLA value from the network meta-analysis, will produce an unfavourable comparison of PVP with OPM, compared with raising the OPLA value to the BKP and PVP network meta-analysis estimate (using the alternate methodology the ICER for BKP/PVP compared with OPM would be that reported in Scenario 5).

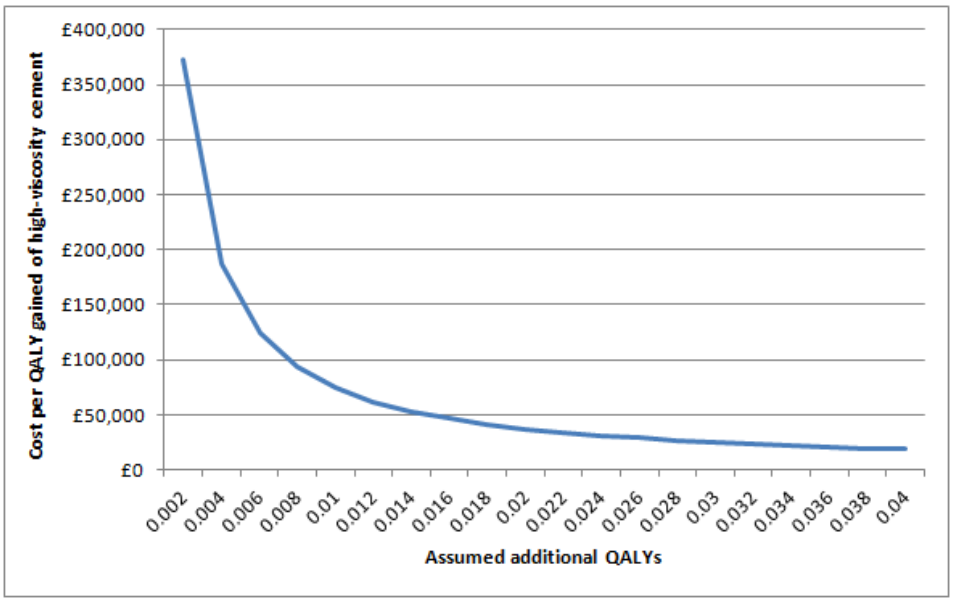
Thus the intervention that is estimated to be most cost-effective is heavily dependent on the assumptions chosen. Given the uncertainty regarding the mortality effects of the treatments (including OPLA) a definitive conclusion cannot be provided. However, given that a facet joint injection is commonly used, relatively inexpensive and may have considerable benefit in up to one third of patients[83] it is likely that this is considered an appropriate first measure.

Additional Exploratory Analyses

The use of high viscosity cement for all patients

The definition of PVP within the analyses is that of low viscosity cement use for the majority of patients and with high viscosity cement being used in the estimated 15% of patients in which our clinical advisor (DW) believed that this was a clinical necessity. An alternative strategy (and the one used by Johnson and Johnson) is to assume that all patients receive Johnson and Johnson's high viscosity cement at an additional cost of £746 per patient. Figure 68 undertakes an exploratory analysis regarding the Cost per QALY gained of using high viscosity cement when the assumed QALY increases associated with the use of high- (rather than) low-viscosity cement is used.

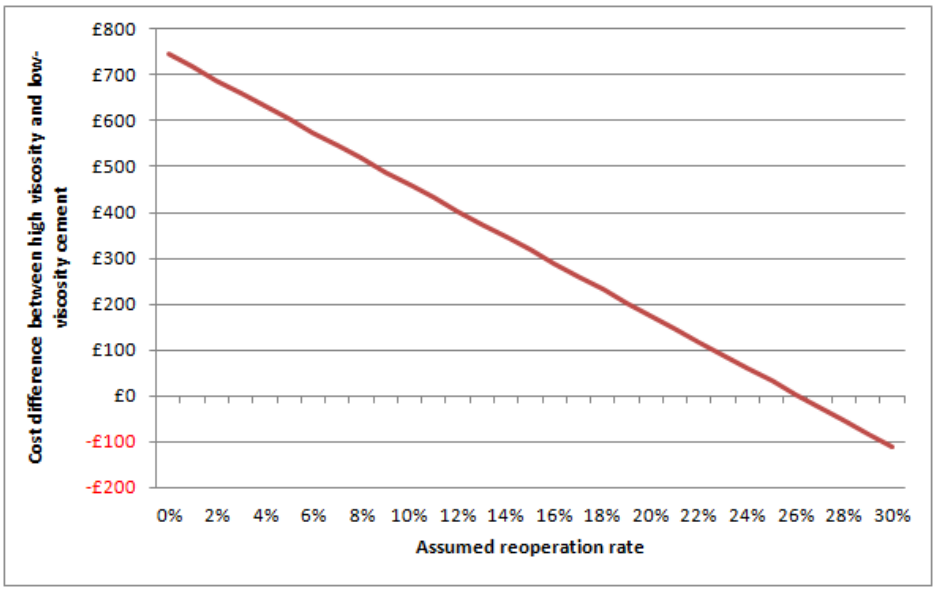
Figure 68: An exploratory analysis of effect of assuming additional QALY gains associated with high-viscosity cement compared with low-viscosity cement.



It is calculated that there would need to be an additional 0.037 QALYs for the cost per QALY gained to be equal to £20,000 per QALY gained. It is noted that this value is greater than the value of 0.02 discounted QALYs assumed in the sensitivity analyses that was estimated assuming that 1 in 1000 people died (incurring a loss of 10 discounted QALYs) and that 1 in 100 people experienced morbidity resulting in a loss of 1 discounted QALYs when using low viscosity cement. As such, it is unlikely that the ICER of high-viscosity cement compared with low viscosity cement would be lower than £20,000 per QALY gained.

The above analysis assumes that costs would remain constant, whereas there is a possibility that operations would need to be re-performed were there to be a problem with low viscosity cement. In order to explore this impact, the costs per QALY gained of high viscosity cement at different levels of re-operation rates were estimated. In order for the graph to be shown the operation costs associated with PVP of £1479 was assumed to be correct, rather than the AIC value used by the Assessment Group. It was assumed that all re-operations would be undertaken using high viscosity cements.

Figure 69: An exploratory analysis of effect of assuming additional re-operations associated with high-viscosity cement compared with low-viscosity cement.



It is calculated that there would need to be a re-operation rate in excess of 25% in order for a strategy of using high viscosity cement in all patients to be cheaper than using it in a selected 15% of patients, assuming that QALYs remained unaltered. Such values have not been reported, with the only identified estimate being less than 1.5%[244] and it not being certain that high-viscosity cement would have prevented the re-operation in each case.

However, it is likely that were re-operations required then there would also be a QALY effect, and thus the cost per QALY gained of selected combinations has been calculated as shown in Table 99.

Table 99 The effect on the cost per QALY gained of high viscosity cement when changing both the QALY gained and the level of re-operations.

		Assumed Re-operation rate				
		0%	**5%**	**10%**	**15%**	**20%**
Assumed QALY increase associated with high-viscosity cement	0.000	Dominated	Dominated	Dominated	Dominated	Dominated
	0.005	£149,200	£120,630	£92,060	£63,490	£34,920
	0.010	£74,600	£60,315	£46,030	£31,745	£17,460
	0.015	£49,733	£40,210	£30,687	£21,163	£11,640
	0.020	£37,300	£30,158	£23,015	£15,873	£8,730
	0.025	£29,840	£24,126	£18,412	£12,698	£6,984
	0.030	£24,867	£20,105	£15,343	£10,582	£5,820

Based on the results shown in Table 99 it is unlikely that a strategy of using high viscosity cement in all patients rather than a subset selected by the clinican would have an ICER less than £20,000 per QALY gained.

Facet Joint Injections

Consideration was given to explicitly modelling the cost-effectiveness of a pathway involving an initial facet joint injection. Our clinical advisor indicates that a facet joint injection is an outpatient procedure, requiring 15 to 20 minutes of fluoroscopy room time, a radiologist and radiographer for this period and incurs approximately £60 of drugs and consumables per case. Our clinical advisor (DW) indicated that the cost of a facet joint injection was unlikely to exceed £200 per patient. As such it is significantly cheaper than PVP or BKP and has been shown both to reduce the numbers of patients progressing to vertebral augmentation and improve the response rate of those requiring PVP.[83] It is currently unclear whether the increased response is due to removing patients who would have healed naturally without augmentation or whether there is a placebo response to the injection. As anecdotally the use of initial facet joint injections appears to be widespread in the UK, and as facet joint injections were neither interventions nor comparators in the NICE evaluation (http://www.nice.org.uk/nicemedia/live/13445/57319/57319.pdf) it was decided to not model facet joint injections.

However, if the use of facet joint injections increases the likelihood of patients responding to vertebral augmentation, and facet joint injection experienced patients were excluded from the RCTs then the benefit of PVP or BKP may be underestimated. An exploratory analysis of the effect of prior facet joint injection has been undertaken. If it is assumed that one third of

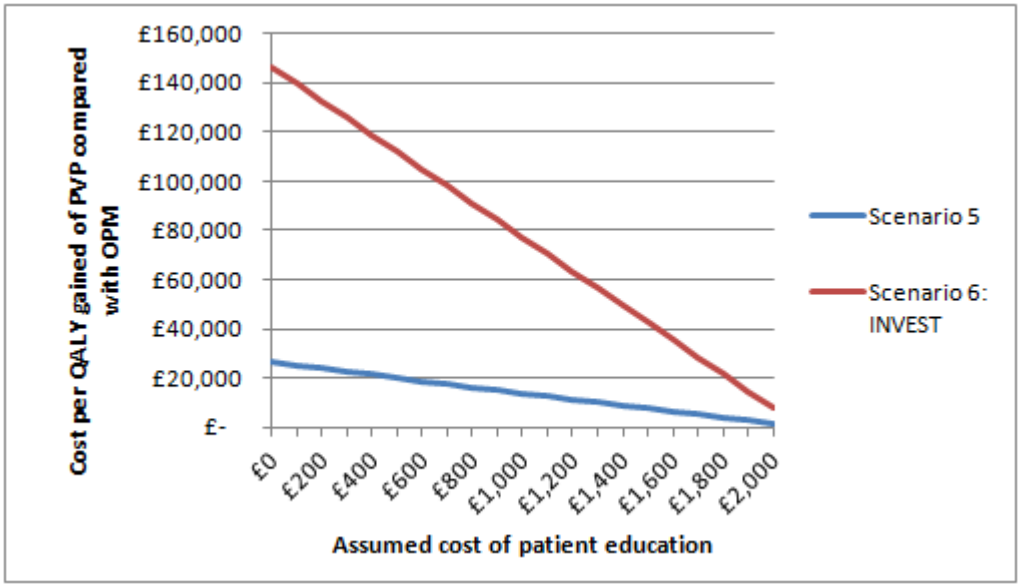
patients would respond to a facet joint injection[83] and that these would have exhibited identical VAS / EQ-5D effects regardless of the treatment arm, then the average VAS or EQ-5D difference shown in the entire population would be estimated to be increased by 50% when just considering those who did not respond to the facet joint injection. If it is assumed that the entire QALY difference is due to VAS/EQ-5D scores (rather than adverse events) then the ICER would be reduced by one third implying that ICERs of £30,000 may be reduced to £20,000 if all patients had a facet joint injection initially.

Additional Patient Education

The double-blinded trials showed minimal difference between PVP and OPLA indicating that the benefit seen within the unblended trials could be driven by a placebo response. It is unclear whether the placebo can be generated by less intensive methods than preparing a patient for an operation but performing OPLA rather than vertebral augmentation. An exploratory analysis has been performed to estimate the maximum expenditure that could be provided in educating patients in order that the OPLA responses were assumed to be generated in people receiving OPM, whilst maintaining a cost per QALY gained ratio of PVP compared to OPM of greater than £20,000.

The analysis assumed that there was no beneficial effect of either PVP or BKP, as in this instance the ICERs for vertebral augmentation were generally low compared to OPLA. The exploratory analysis evaluated three scenarios: Scenario 5; Scenario 6: Buchbinder and Scenario 6: INVEST as these are the studies that explicitly take the relationship between OPLA and PVP into account. It was assumed that OPM would have identical results to those for OPLA, and that ignoring patient education costs, OPM would cost £2111 less than PVP comprising of a cost of £800 for the PVP equipment and £1311 associated with the operation. As the results from Buchbinder et al produced identical improvements from PVP and OPLA then for this study the decision simplified into cost-minimisation (excluding the possibility of adverse events) and OPM would be seen to dominate PVP if the education costs were below £2111 per person and be dominated if the cost was above this value. For Scenario 5 and Scenario 6: INVEST Figure 70 indicates the cost per QALY given different assumed costs of patient education. It is commented that this analysis is exploratory. It is not known whether the OPLA placebo response could be generated by education nor is the likely costs of patient education known.

Figure 70: An exploratory analysis of the cost per QALY gained of PVP versus OPM assuming that the response of OPLA could be generated in OPM patients through education.



7. ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES

Place of PVP and BKP in the treatment pathway

BKP and PVP are remedial measures, and prevention of VCFs should ideally be pursued to avoid unnecessary surgical procedures. There is evidence of suboptimal utilisation of pharmacological treatments, such as bisphosphonates, for at-risk patients.[365,366] Proactive case selection strategies have shown promise for enhancing appropriate prescribing.[366]

Care providers will need to consider at what point in the treatment pathway PVP or BKP should be offered. Wilson and colleagues have suggested that PVP and BKP should not be considered as a first line of treatment.[367] The same authors have found evidence to suggest that facet joint injections may be clinically effective in a substantial minority of patients.[83] Hence, these investigators recommended an initial period of conservative management, followed by a facet joint injection, with augmentation offered only after failure of both these less invasive approaches.

Similarly, previous NICE guidance from 2003 suggests PVP should only be used if pain is refractory to nonsurgical pain management (http://www.nice.org.uk/nicemedia/pdf/IPG12A4Updated.pdf), while NHS Oxfordshire guidance (http://www.oxfordshirepct.nhs.uk/professional-resources/priority-setting/lavender-statements/documents/LS154_Vertebroplasty.pdf) suggests PVP should be offered only after at least 4 weeks' OPM, including local anaesthetic / steroid injection to affected area.

Ethical issues and the placebo response

Findings from the existing literature suggest that vertebral augmentation is substantially better than conservative management, but it is uncertain whether it is more beneficial than OPLA. It has been suggested that the lack of evidence of a demonstrable benefit over OPLA represents a powerful placebo response to PVP, due to factors such as the positive expectations of patients and clinicians, and activation of pain-reducing neurobiological pathways.[333]

This raises important issues with respect to medical ethics. If the positive effect of PVP is unrelated to the injection of cement, one would have to ask whether the benefits outweigh the known risks of the procedure. Miller et al.[333] have argued that it is not necessarily unethical to provide a minimally invasive procedure with the aim of generating a powerful placebo response. They note a growing body of research showing that the

placebo effect is associated with real neurophysiological effects that may lead to clinically meaningful improvements.[368-370] With respect to pain, for example, there is evidence that the 'placebo response' involves the activation of endogenous opioids and dopamine pathways – that is, it has specific mechanisms for efficacy.[371,372] There has been some evidence provided that the use of an initial facet joint injection may produce clear benefits in around a third of patients, with a reduced risk profile and low cost and this should be considered as an initial first response.

The registry data showing improved survival rates following vertebral augmentation[174] further complicates the ethical issues related to PVP and BKP. The implications of these findings hinge to a substantial degree on whether the improvement was due to biomechanical factors directly associated with the injection of cement (e.g. correction of kyphotic wedge angle and vertebral body height). On the other hand, there may have been unobserved selection factors for the procedure which were directly related to mortality, and there were a number of methodological issues with the registry findings, as discussed in Appendix 12.

8. DISCUSSION

8.1 Statement of principal findings

In unblinded trials, PVP and BKP perform significantly better than OPM in improving HRQoL, functional ability, and pain in the short- to medium-term. However, there is no convincing evidence that vertebral augmentation provides any substantial benefits above OPLA. In addition, OPLA is not associated with the serious adverse events that can result from vertebral augmentation. The two double-blind, placebo-controlled trials[101,102] provide the highest level of clinical effectiveness evidence to date, although these studies were not large and may have patient selection issues. The ongoing VERTOS IV trial[373] will provide important additional evidence on the PVP versus OPLA comparison.

[redacted], and pooled 12-month mortality rates from three RCTs[17,146,185] are slightly suggestive that vertebral augmentation may have mortality benefit. However, there were limitations associated with these analyses to the extent that no definitive statement on the presence and size of any mortality benefit could be made. The potential presence of a mortality benefit is a key issue for both the clinical and cost-effectiveness of vertebral augmentation. If differential benefits for PVP and BKP (favouring BKP) are assumed, then BKP is more efficacious and is associated with a cost per QALY gained value of below £20,000. If equivalent gains for both PVP and BKP are assumed then PVP is estimated to dominate BKP and to have an ICER below £10,000 per QALY gained in all bar one scenario where OPLA was assumed to have an identical effect on mortality. If BKP and PVP are assumed to have no mortality benefit, PVP dominated BKP, although the ICER compared with OPLA and OPM was dependent on the assumptions made. If data from the two blinded studies were used then OPLA was estimated to be the most cost-effective and if this was not considered an appropriate comparator the ICER between PVP and OPM was estimated to range from £15,000 to £40,000 per QALY. The analyses for PVP were assumed using low-viscosity cement for the majority of patients with selected patients receiving high-viscosity cements. Exploratory analyses assuming that all patients received high-viscosity cement indicate that the cost-effectiveness of this strategy was likely to be greater than £20,000 per QALY gained.

8.2 Strengths and limitations of the assessment

To our knowledge, this is the first review to attempt an aggregation of clinical and cost effectiveness from clinical trials of PVP and BKP for osteoporotic VCFs. The robustness of the review process was enhanced by a comprehensive search strategy, including a broad search of databases, contact with clinical experts, and manual searches of the bibliographies of retrieved studies. The robustness of the findings was also enhanced by independent data extraction, assessment of quality, and study inclusion, by two reviewers. The assessment of clinical effectiveness included RCTs only, while the assessment of safety included data from RCTs, large case series, and individual case reports. The analyses conducted the most robust mapping of VAS to EQ-5D of which we are aware, and undertook a network meta-analysis on the VAS data. Extensive scenario and sensitivity analyses were conducted to explore a wide range of different assumptions. Insufficient evidence, particularly on the impact of BKP, PVP and OPLA on mortality rates, means that no definitive conclusion can be made.

However, a systematic review can only be as good as the studies it includes. With respect to the data set, the most serious methodological problem was the lack of blinding in all studies except INVEST and Buchbinder. Unblinding in surgical studies has been linked to a 25% over-estimation of treatment effect.[337] As Buchbinder and Kallmes have pointed out, the improvement in the treatment groups of the blinded trials was not dissimilar to those seen in the treatment groups of the unblinded trials.[374] The assessment of BKP effectiveness was particularly limited, since only one open-label RCT comparing BKP with non-surgical management was available, while the only study to compare BKP with PVP showed a number of potential sources of bias.

A further limitation of these findings was the use of pain as a primary outcome. As others have argued, pain measurement may be confounded by a number of factors, including pain threshold, analgesia, and level of activity.[83] Back pain-related disability and quality of life may provide more objective and clinically meaningful measures.[83] However, these outcomes were measured in heterogeneous ways among the trials, precluding statistical aggregation of the data. Measures of vertebral body height and angular deformity may also be more useful clinical outcomes than pain, insofar as improvements could enhance mobility and stave off deterioration of cardio-pulmonary function.[345] Four studies (Blasco, Farrokhi, FREE, Liu) reported these outcomes, but it was not possible to aggregate their findings due to the heterogeneous approaches that were taken to wedge angle and vertebral body height measurement.

This review was specific to the population of people with painful osteoporotic VCFs. Hence, the results are not necessarily generalisable to VCFs of other origins (e.g. multiple myeloma, traumatic). Discussions of generalisability among the studies were usually cursory. For example, several studies did not present data on the ethnic composition of their samples, nor did they comment on the implications of this for generalisability. On the other hand, the age and gender makeup of the study samples was fairly representative of the wider osteoporotic population in the UK. A higher proportion of females took part in the trials (typically around 70%); and the mean sample age was usually early to mid 70s. In addition, since all studies, with the exceptions of INVEST and FREE, were carried out exclusively outside the UK, the generalisability of the findings to the UK population of people with painful osteoporotic VCFs is unclear.

Perhaps most importantly, we were unable to establish whether percutaneous vertebral augmentation leads to changes in rates of mortality. [redacted]

However, due to lack of data on causes of death and other confounding factors, causal mechanisms other than the augmentation procedures cannot be ruled out at this stage. Data on 12-month mortality from three RCTs[17,146,185] were pooled in this review. Although the point estimate slightly favoured PVP, it was not possible to rule out no effect. More problematically, as noted by Aebi,[21] 12 months is unlikely to be long enough to capture longer-term implications of kyphotic deformity and impaired cardiopulmonary function associated with VCFs.

8.3 Uncertainties

The key uncertainty is whether vertebral augmentation provides a mortality effect over OPM or OPLA. A definitive causal relationship cannot be inferred from the available observational data, and it would be difficult to conduct RCTs with adequate power and follow-up duration to fully explore this. While it seems likely that PVP is no more effective than OPLA in improving functional ability, pain, and QoL, there is yet to be a head to head comparison of BKP and OPLA, although evidence from the network meta analysis indicates that this too would not be expected to be more effective than OPLA in improving functional ability, pain, and QoL. It is also not known if there may be ways to generate the apparently high placebo response without resorting to cement injection or OPLA, although an exploratory analysis has been conducted to indicate how much could be spent on patient education with the cost per QALY of PVP remaining above £20,000 per QALY gained compared with education. This value was seen to at least £500 per patient and could be considerably more.

Finally, further evidence for the effect of vertebral augmentation on restoration of vertebral body height and sagittal balance is required.

8.4 Other relevant factors

Risks to staff

There has been some discussion over the past decade concerning the risk to staff performing vertebroplasty and kyphoplasty procedures of radiation exposure: this risk is low, but of potential importance. Fitousi et al. estimated that vertebroplasty practitioners could perform 150 vertebroplasty procedures annually without exceeding annual dose constraints, while Harstall[375] estimated an annual risk of 0.0025% for fatal cancer of the thyroid, and a small to medium risk of developing any cancer of 0.025%. However, these risks can be somewhat mitigated by following a number of precautionary measures[376] including the use of protective lead gloves,[377] other shielding techniques,[378] and the use of a combined CT – fluoroscopy approach to imaging, as opposed to fluoroscopy only.[379]

In addition, some staff have experienced an idiosyncratic reaction or asthma exacerbation in response to PMMA vapour even though exposure during a typical PVP case or list is below the established occupational exposure limits for staff working with PMMA.[380]

9. CONCLUSIONS

9.1 Implications for service provision

PVP is likely to provide greater clinical benefits than conservative management, and may be one way to mitigate some of the problems associated with the latter approach. However, two blinded randomised controlled trials indicate that PVP does not appear to be any more effective than administration of local anaesthesia to the affected area in improving pain, quality of life, or back-related functional ability. As yet, there are no well-designed, double-blind, placebo controlled trials of BKP. Hence, although this procedure is likely to be beneficial in comparison to conservative treatment, its effectiveness compared with local anaesthetic was estimated through a network meta-analysis, which indicated that PVP and BKP had similar long-term VAS scores. Although some data suggest that PVP and BKP may lead to long-term reductions in mortality, it is not yet clear whether this effect is due to a specific mechanism of the procedures, or to other extraneous factors.

The cost-effectiveness of each intervention is strongly influenced by the assumed mortality benefit. As the evidence for mortality benefits for PVP and BKP is limited it is unclear whether these interventions will be recommended and hence the implications for service provision are unknown.

9.2 Suggested research priorities

The effect of local anaesthesia on functional and pain-related improvements in people with VCFs remains a contentious issue. Buchbinder and Kallmes have argued that injecting a short-acting anaesthetic over the pedicle of the fractured vertebral body (as per the INVEST study) would be unlikely to provide sustained benefits.[374] Others have argued that anaesthesia injection may have specific mechanisms of efficacy,[111,202] and there is some limited evidence to support this proposition. For example, Riew et al. conducted a RCT of the efficacy of selective nerve root blocks in the treatment of lumbar radiculopathy, and found that local anaesthetic showed an effect long after its expected duration.[113] More recently, Wilson et al. found that administering local anaesthetic with steroid facet joint injection to the most painful level led to substantial improvements in approximately one third of a cohort of 75 patients with painful VCFs.[83] However, the facet joint block may provide additive benefits to those provided by anaesthesia alone. A study comparing anaesthesia and facet joint injection with anaesthesia only, would be useful to explore any possible placebo effects in these approaches.

There are few clinical trials of vertebral augmentation with proper blinding and placebo controls. The total number of patients in the Buchbinder and INVEST studies was low, even when aggregated (n=209). Further double-blinded, placebo-controlled RCTs would be helpful in confirming the findings of these trials. In addition, there are currently no double-blinded, placebo-controlled RCT data on BKP.

There is ambiguity regarding patient selection for PVP and BKP. Critical commentaries on the Buchbinder and INVEST trials have suggested that only patients with <6 weeks of pain should be treated.[106,334,347] However, a division of patients in the Buchbinder and INVEST trials into <=6 weeks of pain and >6 weeks of pain indicated very similar adjusted between group differences indicating that this theory was incorrect.[109] As pain tends to spontaneously improve during the acute phase, others have suggested that, in order to avoid unnecessary surgical interventions, PVP and BKP should only be used in cases of intractable and long-term pain.[336] To this end, more IPD analyses comparing effectiveness in acute and long-term pain from blinded trials would be helpful.

Both spinal deformity and sagittal balance are both important measures of VCF severity. If balance becomes unstable then this may be a further pain generator alongside micro-movement of fractures and pain on adjacent joints. Moreover, increasing kyphosis and reduction of vertebral body height can lead to deterioration in cardiopulmonary function, and ultimately, death. However, the effectiveness of PVP and BKP in restoring these morphometric parameters is yet to be studied in high quality trials.

Finally, the effect of vertebral augmentation on mortality is a potentially important area for further exploration as this has been shown to strongly influence the cost-effectiveness of the interventions. Data from the US and Germany suggested that PVP and BKP were associated with reduced mortality rates compared with people with VCFs who did not undergo cement augmentation.[381] However, formal analyses of the data provided to the Assessment Group highlighted potential methodological limitations meaning that a definitive conclusion could not be formed. Notwithstanding the various methodological and ethical issues which must be carefully addressed in RCTs of surgical procedures,[382] it would be desirable to have additional double-blinded RCTs of vertebral augmentation with adequate power and follow-up length to investigate this possible effect further. Alternatively, a prospective observational database containing as many co-variates as were feasible could provide beneficial data on the likely mortality impacts.

10. APPENDICES

Appendix 1: Literature Search Strategies

Medline clinical effectiveness search strategy (Ovid)

1. Vertebroplasty/
2. Kyphoplasty/
3. vertebroplasty.ti,ab.
4. kyphoplasty.ti,ab.
5. bone void fill*.ti,ab.
6. injectable bone cement*.ti,ab.
7. osteoplastic procedure*.ti,ab.
8. vertebral* augmentation*.ti,ab.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8

Medline Economics Strategy (SIGN Strategy)

1 Economics/

2 "costs and cost analysis"/

3 Cost allocation/

4 Cost-benefit analysis/

5 Cost control/

6 Cost savings/

7 Cost of illness/

8 Cost sharing/

9 "deductibles and coinsurance"/

10 Medical savings accounts/

11 Health care costs/

12 Direct service costs/

13 Drug costs/

14 Employer health costs/

15 Hospital costs/

16 Health expenditures/

17 Capital expenditures/

18 Value of life/

19 Exp economics, hospital/

20 Exp economics, medical/

21 Economics, nursing/

22 Economics, pharmaceutical/

23 Exp "fees and charges"/

24 Exp budgets/

25 (low adj cost).mp.

26 (high adj cost).mp.

27 (health?care adj cost$).mp.

28 (fiscal or funding or financial or finance).tw.

29 (cost adj estimate$).mp.

30 (cost adj variable).mp.

31 (unit adj cost$).mp.

32 (economic$ or pharmacoeconomic$ or price$ or pricing).tw.

33 Or/1-32

Appendix 2: Data extraction form

Randomised controlled trials data extraction form

(based on NHS CRD Report No. 4. [NHS Centre for reviews and Dissemination. *Report 4: Undertaking systematic reviews of research on effectiveness; CRD's guidance for those carrying out or commissioning reviews.* York: University of York; 2001.])

STUDY & DESIGN	DATA EXTRACTION	
Trial	REVIEW DETAILS	
	Author, year	
Study design:	**Objective**	
	Publication type (ie full report or abstract)	
	Country of corresponding author	
	Language of publication	
	Sources of funding	
	INTERVENTIONS	
	Focus of interventions (comparisons)	
	Description	
	T1: Intervention group	
	T2: Control group	
	Intervention site (health care setting, country)	
	Procedure performed by	
	Length of follow up	
	STUDY CHARACTERISTICS	

STUDY & DESIGN	DATA EXTRACTION	
	Method of randomisation	
	Description	
	Generation of allocation sequences	
	Allocation concealment?	
	Blinding level	
	Numbers included in the study	
	Numbers randomised	
	POPULATION CHARACTERISTICS	
	Target population (describe)	
	Inclusion / exclusion criteria (n)	
	Recruitment procedures used (participation rates if available)	
	Characteristics of participants at baseline	
	Age (mean yr.)	
	Female	
	Median duration of back pain in weeks (IQR)	
	Duration of symptoms <6 weeks	
	BMI	
	Median duration of corticosteroid use in years (IQR)	

STUDY & DESIGN	DATA EXTRACTION	
	Pain score (scale of 0-10) Overall At rest In bed at night	
	QUALEFFO total score	
	AQoL score	
	RDQ score	
	EQ-5D score	
	Timed Up and Go test (Seconds)	
	Medication for osteoporosis Any Calcium supplements Vitamin D Bisphosphonates	
	One or more previous vertebral fractures	
	Opioids for pain	
	BMD T-score $\leq$2.5 Lumbar Femoral neck	
	Severity of fracture (no/total no of fractures) assessed by Genant's semiquantitative system Mild Moderate Severe	

STUDY & DESIGN	DATA EXTRACTION	
	No of vertebral bodies treated 1 2	
	Other information	
	Were intervention and control groups comparable?	

	OUTCOMES	
	Definition of primary outcomes	
	Definition of secondary outcomes	
	Definition of tertiary outcomes	
	Definition of other outcomes	
	ANALYSIS	
	Statistical techniques used	
	Intention to treat analysis	
	Does technique adjust for confounding?	
	Power calculation (priori sample calculation)	
	Attrition rates (overall rates) i.e. Loss to follow-up	
	Was attrition adequately dealt with?	
	Number (%) followed-up from each condition	
	RESULTS	

	Adverse events	
	Other information	
	SUMMARY	
	Authors' overall conclusions	
	Reviewers comments	

Appendix 3: Original quality assessment checklist (adapted from Ploeg et al 2006[87])

Criterion	**Yes**	**No**	**Unclear**	**Not applicable**
Is a control group present? If yes:				
Was a method of randomisation performed?				
Was the treatment allocation concealed?				
Were co-interventions avoided or comparable?				
Was the outcome assessor blinded to the intervention?				
Were the outcome measures relevant?				
Was the withdrawal/drop-out rate described and acceptable?				
Was the timing of the outcome assessment comparable in both groups?				
Did the analysis include an intention-to-treat analysis?				
Were the eligibility criteria specified?				
Were the groups similar at baseline regarding the most important prognostic indicators?				
Were the index and control interventions explicitly described?				
Were adverse effects described?				
Was a short-term follow-up measurement performed?				
Was a long-term follow-up measurement performed?				
Was the sample size for each group described?				
Were point estimates presented for the primary outcome measures?				
Were measures of variability presented for the primary outcome measures?				
Was a valid questionnaire, eg concerning pain and quality of life, used?				

Appendix 4: Revised quality assessment checklist

Criterion		**Yes**	**No**	**Unclear**	**Not applicable**
Internal validity					
Selection bias	Was the method used to assign participants to treatment groups really random? (see Note A)				
	What method of assignment was used?				
	Was the allocation of treatment concealed? (see Note B)				
	What method was used to conceal treatment allocation?				
Performance bias	Were co-interventions avoided or comparable?				
	Were the participants who received the intervention blinded to the treatment allocation?				
	Was the success of the blinding procedure assessed?				
Detection bias	Were the outcome assessors blinded to the treatment allocations?				
	Was the success of the blinding procedure assessed?				
	Was the timing of the outcome assessment comparable in both groups?				
Attrition bias	Were at least 80% of the participants originally randomised to treatment followed up in the final analysis?				
	Were the reasons for withdrawal stated?				
	Was the number of participants randomised to each group stated?				
	Did the report state the number of participants in each group who were				

	included in the final analysis?				
	Were there any unexpected imbalances in drop-outs between groups?				
	If there were unexpected imbalances, were they explained or adjusted for?				
	Was an intention-to-treat analysis included?				
	If an intention-to-treat analysis was included, was it appropriate, and were appropriate methods used to account for missing data?				
Reporting bias	Is there any evidence to suggest that the authors measured more outcomes than they reported?				
Other bias	Were the groups similar at baseline in terms of the most important prognostic indicators?				
External validity (generalisability)					
Were the eligibility criteria for study entry specified?					
Were the index and control interventions explicitly described?					
Were the skills, training, and experience of the operator described?					
Were the outcome measures relevant?					
Was a valid instrument used to measure each outcome?					
Was a short-term follow-up measurement performed? (<3 months after randomisation)					
Was a long-term follow-up measurement performed? (≥12 months after randomisation)					
Were adverse effects adequately described?					
Precision					
Was the study powered to detect differences in outcome?					
Were point estimates presented for the primary outcome measures?					

Were measures of variability presented for the primary outcome measures?				

Note A: acceptable methods of randomisation include:

- using a random number table
- using a computer random number generator
- tossing a coin
- shuffling cards or envelopes
- throwing dice
- drawing lots
- minimisation

Unacceptable methods include:

- a process which includes the patient's date of birth, date of admission, hospital or clinic record number
- clinician judgment
- patient preference
- laboratory test results
- availability of the intervention.

Note B: acceptable methods of allocation concealment include:

- central allocation (including telephone, web-based, and pharmacy-controlled randomisation)
- sequentially numbered opaque sealed envelopes

Unacceptable methods include:

- an open random allocation schedule (eg a list of random numbers)
- assignment envelopes used without appropriate safeguards
- alternation
- date of birth
- case record number
- or any other explicitly unconcealed procedure

Appendix 5: Details of studies which were potentially relevant to the review of clinical effectiveness copies of which could not be obtained within the study timescale

39th Annual Meeting of the Spanish Society of Neuroradiology SENR - 6th Congress of the Portuguese Society of Neuroradiology, SPNR. Neuroradiology Conference.

Baier, M., et al. “Pain reduction and verbal redressement by kyphoplasty”. Osteologie 16.3 (2007): 173-175

Bobra, S., et al. "Early outcomes in osteoporotic patients with painful vertebral body fractures treated with percutaneous vertebroplasty." Radiology 221 (2001): 136.

Carlier, R. Y., et al. "Percutaneous vertebroplasty and local kyphosis correction." Radiology 225.2 (2002): 514.

Hoffmeister, E. "Balloon kyphoplasty: continuing evidence of efficacy in treating vertebral collapse and fracture." Bone & Joint 13.6 (2007): 61-65.

Kasperk, C. "Lkypho-vertebroplasty and non pharamcologic treatment." Annals of the Rheumatic Diseases 65 (2006): 14.

Kim, A. K., et al. "Modified transpedicular approach for percutaneous vertebroplasty: Holo-vertebral body filling using a single injection." Radiology 217 (2000): 510.

Kobayashi, T., T. Takanaka, and O. Matsui. "Percutaneous vertebroplasty-guided by CT fluoroscopy." Radiology 217 (2000): 527.

Kraus, J., W. Achatz, and H. G. Gorzer. "Pelvic and crural phlebothrombosis as complication of percutaneous vertebroplasty." Rofo-Fortschritte Auf dem Gebiet der Rontgenstrahlen und der Bildgebenden Verfahren 175.4 (2003): 565-66.

Mallampati, G. K., et al. "Functional outcome and pain modification following vertebroplasty." Radiology 225 (2002): 614.

Oka, M. and P. A. Westesson. "Vertebroplasty can improve pain and mobility." Radiology 225 (2002): 513.

Ruefenacht, D. A., et al. "Vertebroplasty: Clinical results and follow-up." Radiology 213P (1999): 416.

Sehgal, M., L. A. Gilula, and D. B. Brown. "Vertebroplasty in patients with symptoms for greater than one year in duration." Radiology 225 (2002): 513-14.

Theodorou, D. J., et al. "Percutaneous balloon kyphoplasty: A novel technique for reducing pain and spinal deformity associated with osteoporotic vertebral compression fractures." Radiology 217 (2000): 511.

Wang, G. H., et al. "[Percutaneous vertebroplasty and conservative therapy for osteoporotic vertebral compression fractures: a clinical comparative study]." Journal of Interventional Radiology 17.9 (2008): 663-67.

Westesson, P. A. and Y. Numaguchi. "Vertebroplasty: Physical examination, plain film, and bone scan can be misleading in preoperative evaluation." Radiology 221 (2001): 618.

Westesson, P. A. and Y. Numaguchi. "Vertebroplasty: Subsequent compression fracture is a common reason for recurrent pain." Radiology 221 (2001): 617.

Appendix 6: Details of included studies relating to trials which met the inclusion criteria for the review of clinical effectiveness

Asterisks indicate the major publications for the study

Blasco, 2012

Blasco J, Garcia A, Manzanera LSR, MacHo JM, Peris P, Jaume P, et al. Randomized trial comparing vertebroplasty and conservative treatment analyzing pain relief and quality of life on the long term basis. CardioVascular and Interventional Radiology 2010 Sep;33(Suppl 2):182-3.

*Blasco JA, Martinez-Ferrer A, Macho Fernández J, San Roman Manzanera L, Pomés Talló J, Carrasco Jordan JLl, et al. Effect of vertebroplasty on pain relief, quality of life and the incidence of new vertebral fractures. A 12-month randomised follow-up, controlled trial. Journal of Bone and Mineral Research 2012;Accepted manuscript online.

Martinez-Ferrer A, Blasco J, Carrasco JL, Monegal A, Pomes J, Guaabens N, et al. Effect of vertebroplasty on the quality of life of patients with pain related to osteoporotic vertebral fractures preliminary results of a randomized trial. Bone 2011 May 7;48(Suppl 2):S161.

Buchbinder, 2009

Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedt CJ, et al. Efficacy and safety of vertebroplasty for treatment of painful osteoporotic vertebral fractures: a randomised controlled trial [ACTRN012605000079640]. BMC Musculoskeletal Disorders 2008;9:156.

*Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedt C, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. New England Journal of Medicine 2009;361(6):557-68.

Staples MP, Kallmes DF, Comstock BA, Jarvik JG, Osborne RH, Buchbinder R. Effectiveness of vertebroplasty using individual patient data from two randomised placebo controlled trials: meta-analysis. BMJ 2011;343:d3952.

Farrokhi, 2011

*Farrokhi MR, Alibai E, Maghami Z. Randomized controlled trial of percutaneous vertebroplasty versus optimal medical management for the relief of pain and disability in acute osteoporotic vertebral compression fractures. Journal of Neurosurgery Spine 2011 May;14(5):561-9.

FREE

Bastian L, van MJ, Boonen S, Ramstam J, Cummings S, Wardlaw D. [1-year results of a randomised, controlled, international, multi-centre study to compare balloon kyphoplasty and non-surgical care of acute compression fractures of vertebral bodies]. Medizinische Klinik 2008;103(3):16.

Boonen S, Wardlaw D, Bastian L, Lips P, Van Meirhaghe J, Cummings S. Balloon kyphoplasty and non-surgical management in patients with acute vertebral body compression fractures: A randomized comparative trial. Calcified Tissue International 2007;80(Suppl 1):S33.

Boonen S, Cummings S, Wardlaw D, Eastell R. Impact of balloon kyphoplasty on quality of life and risk of recurrent vertebral fractures: A randomized trial in patients with acute vertebral compression fractures. Calcified Tissue International 2008;82(Suppl 1):S40-S41.

Boonen S, Van MJ, Bastian L, Cummings SR, Ranstam J, Tillman JB, et al. Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. Journal of Bone & Mineral Research 2011 Jul;26(7):1627-37.

Van Meirhaeghe JK, Boonen S, Bastian L, Cummings S, Ranstam J, Tillman J, et al. A randomized trial of balloon kyphoplasty and nonsurgical care for patients with acute vertebral compression fractures: Two year results. Osteoporosis International 2010 Dec;21(Suppl 5):S667-S668.

Wardlaw D, Boonen S, Bastian L, Van Meirhaeghe J, St Jan AZ. An international multicenter randomized comparison of balloon kyphoplasty and nonsurgical care in patients with acute vertebral body compression fractures. Journal of Bone and Mineral Research 2007;22(7):1119.

*Wardlaw D, Cummings SR, Van Meirhaeghe J, Bastian L, Tillman JB, Ranstam J, et al. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet 2009;373:1016-24.

INVEST

Brinjikji W, Comstock BA, Heagerty PJ, Jarvik JG, Kallmes DF. Investigational Vertebroplasty Efficacy and Safety Trial: detailed analysis of blinding efficacy. Radiology 2010 Oct;257(1):219-25.

Gray LA, Jarvik JG, Heagerty PJ, Hollingworth W, Stout L, Comstock BA, et al. INvestigational Vertebroplasty Efficacy and Safety Trial (INVEST): a randomized controlled trial of percutaneous vertebroplasty. BMC Musculoskeletal Disorders 2007;8:126.

*Kallmes DF, Comstock BA, Heagerty PJ, Turner JA, Wilson DJ, Diamond TH, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. New England Journal of Medicine 2009;361(3):569-79.

Kallmes DF, Comstock BA, Gray LA, Heagerty PJ, Hollingworth W, Turner JA, et al. Baseline pain and disability in the Investigational Vertebroplasty Efficacy and Safety Trial. Ajnr: American Journal of Neuroradiology 2009 Jun;30(6):1203-5.

Staples MP, Kallmes DF, Comstock BA, Jarvik JG, Osborne RH, Buchbinder R. Effectiveness of vertebroplasty using individual patient data from two randomised placebo controlled trials: meta-analysis. BMJ 2011;343:d3952.

Liu, 2010

*Liu JT, Liao WJ, Tan WC, Lee JK, Liu CH, Chen YH, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporosis International 2010 Feb;21(2):359-64.

Rousing, 2009

*Rousing R, Andersen MO, Jespersen SM, Thomsen K, Lauritsen J. Percutaneous vertebroplasty compared to conservative treatment in patients with painful acute or subacute

osteoporotic vertebral fractures. Three-months follow-up in a clinical randomized study. Spine 2009;34(13):1349-54.

Rousing R, Hansen KL, Andersen MO, Jespersen SM, Thomsen K, Lauritsen JM. Twelve-months follow-up in forty-nine patients with acute/semiacute osteoporotic vertebral fractures treated conservatively or with percutaneous vertebroplasty. A clinical randomized study. Spine 2010;35(5):478-82.

VERTOS

*Voormolen MHJ, Mali WPTM, Lohle PNM, Fransen H, Lampmann LEH, van der Graaf Y, et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteoporotic vertebral compression fractures. The VERTOS study. American Journal of Neuroradiology 2007;28(3):555-60.

VERTOS II

Klazen CAH, Verhaar HJJ, Lampmann LEH, Juttmann JR, Blonk MC, Jansen FH, et al. VERTOS II: percutaneous vertebroplasty versus conservative therapy in patients with painful osteoporotic vertebral compression fractures; rationale, objectives and design of a multicenter randomized controlled trial. Trials 2007;8(33).

Klazen C, Lohle P, Jansen F, Schoemaker M, Elgersma O, Van EK, et al. 1-year results of the VERTOS II trial: Vertebroplasty versus conservative therapy. CardioVascular and Interventional Radiology 2009 Sep;32(Suppl 2):313.

*Klazen CAH, Lohle PNM, Jansen FH, Tielbeek AV, Blonk MC, Venmans A, et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. Lancet 2010;376:1085-92.

Klazen CAH, Venmans A, de Vries J, van Rooij WJ, Jansen FH, Blonk MC, et al. Percutaneous vertebroplasty is not a risk factor for new osteoporotic compression fractures: results from VERTOS II. American Journal of Neuroradiology 2010;31(8):1447-50.

Venmans A, Klazen CAH, Lohle PNM, van Rooij WJ, Verhaar HJJ, de Vries J, et al. Percutaneous vertebroplasty and pulmonary cement embolism: results from VERTOS II. American Journal of Neuroradiology 2010;31(8):1451-3.

Venmans A, Klazen CA, van Rooij WJ, de Vries J, Mali WP, Lohle PN. Postprocedural CT for perivertebral cement leakage in percutaneous vertebroplasty is not necessary - results from VERTOS II. Neuroradiology 2011;53(1):19-22.

Appendix 7: Table of excluded studies with rationale

This is not intended to be an exhaustive list of every study examining the intervention. It includes studies identified by the electronic searches which initially appeared to be relevant to the systematic review of clinical effectiveness but on closer inspection were not deemed to be relevant and/or valid. In addition, it includes RCTs cited in the submission by Johnson & Johnson[35] which did not meet the review group's inclusion criteria. The submissions by Medtronic and Synthes did not include any such RCTs.

Study name	**Reason for exclusion**
Anselmetti 2008[344]	Cited by Johnson & Johnson. Comparison not relevant (high-viscosity vs low-viscosity cement)
Appel 2001[383]	Includes patients with cancer; results for patients with osteoporosis not reported separately
Baerlocher 2010[384]	Cited by Johnson & Johnson. Discussion paper
Becker 2007[385]	Cited by Johnson & Johnson. Intervention not relevant (prophylactic BKP)
Bian 2006[386]	Not clear that this was limited to patients with osteoporotic fracture
Boonen 2010[387]	Does not include data not found in the included publications relating to the FREE study
Buchbinder 2009[388]	Does not include data not found in the main publication of the study by Buchbinder et al.
Buchbinder 2011[389]	Discussion paper
Buchbinder 2010[374]	Discussion paper
Chen C 2010[350]	Cited by Johnson & Johnson. Comparison not relevant (unipedicular vs bipedicular BKP)
Chen L 2011[351]	Cited by Johnson & Johnson. Comparison not relevant (unipedicular vs bipedicular BKP)
Cummings 2009[390]	Does not include data not found in the included publications relating to the FREE study
Figueiredo 2009[391]	Cited by Johnson & Johnson. Comparison not relevant (traditional vs side-opening cannula for PVP)

Firanescu 2011[373]	Does not include any results
Gray 2009[392]	Cited by Johnson & Johnson. Not a randomised study
Holden 2002[393]	Includes patients with cancer; results for patients with osteoporosis not reported separately
Mao 2007[394]	Cited by Johnson & Johnson. Comparison not relevant (carbonated hydroxyapatite cement vs PMMA)
Ramaswamy 2000[395]	Not randomised: patients divided into groups on the basis of duration of pain
Smith 2009[396]	Does not include data not found in the included publications relating to the FREE study

Appendix 8: Data abstraction tables

Table 100: Technical characteristics of the included studies

Study	Pre-procedural imaging method used to determine fracture characteristics	Background and experience of operator performing vertebroplasty or kyphoplasty	Anaesthetic	Approach for vertebroplasty or kyphoplasty	Type of cement used	Mean number of vertebrae treated in index procedure	Mean (SD) volume of cement injected (per vertebra)
Blasco 2012[146]	MRI for all patients; if MRI inconclusive, bone scan	Experienced neurointerventional radiologists	Not specified	Mostly bilateral transpedicular.	PMMA manufactured by Exolent Spine, Elmdown, London UK	2.46	NR
Buchbinder 2009[101]	MRI (if MRI not feasible, eg because of contraindications, CT scan and bone scan[150])	Experienced interventional radiologists with formal training in vertebroplasty and appropriate certification, who were actively performing the procedure	Neurolept sedation/analgesia using midazolam and fentanyl[150]	Unipedicular. Satisfactory infiltration of the vertebral body was confirmed radiographically. Injection was stopped when substantial resistance was met or when the cement reached the posterior quarter of the vertebral body, or if cement leaked into extraosseous	PMMA: manufacturer and viscosity not specified	1.18	2.8 (1.2) ml

Study	Pre-procedural imaging method used to determine fracture characteristics	Background and experience of operator performing vertebroplasty or kyphoplasty	Anaesthetic	Approach for vertebroplasty or kyphoplasty	Type of cement used	Mean number of vertebrae treated in index procedure	Mean (SD) volume of cement injected (per vertebra)
				structures or veins. A bipedicular approach was used only if there was inadequate instillation of cement with the unipedicular approach.			
Farrokhi [147]	X-ray and MRI for all patients	Neurosurgeon	Conscious sedation (iv fentanyl and midazolam) in 10/40 patients; GA in 30/40 patients	Unilateral parapedicular approach in 35 patients (87.5%); bipedicular approach in 5 patients (12.5%). Cement was injected using fluoroscopic monitoring with a C-arm unit in both planes. A bilateral approach was used only if there fluoroscopy indicated inadequate instillation of cement with the unilateral approach.	PMMA: manufacturer and viscosity not specified	2.5	3.5 ml (median 3.1 ml, range 1-5.5 ml) (This figure applies only to patients with a single treated fracture; for patients with several treated fractures, a total figure is given)
FREE[151]	MRI for all patients	Not specified	Most procedures were done under GA, but 6/149 patients	Bilateral, transpedicular, or extrapedicular.	PMMA manufactured by	1.3	NR

Study	Pre-procedural imaging method used to determine fracture characteristics	Background and experience of operator performing vertebroplasty or kyphoplasty	Anaesthetic	Approach for vertebroplasty or kyphoplasty	Type of cement used	Mean number of vertebrae treated in index procedure	Mean (SD) volume of cement injected (per vertebra)
			had conscious or deep sedation with local anaesthesia		Medtronic Spine LLC. Viscosity not specified.		
INVEST[102]	Plain film x-ray or MRI;[152] requirement, for fractures of uncertain age as indicated by pain onset, of marrow oedema on MRI or increased vertebral-body uptake on bone scan	Highly experienced practitioners (discipline not specified, but said by Orr[202] to be interventional radiologists) who had performed a mean of approximately 250 procedures (range 50-800).	Conscious or deep sedation according to the treating physician's usual practice[335]	Typically unipedicular.[199] PMMA infused under constant lateral fluoroscopy into the vertebral body, and iInfusion was stopped when the cement reached to the posterior aspect of the vertebral body or entered an extraosseous space such as the intervertebral disk or an epidural or paravertebral vein.	Barium-opacified PMMA; manufacturer and viscosity not specified	1.4	NR
Liu 2010[148]	Not specified	Not specified	Intravenous GA (Propofol) plus 2% xylocaine injected locally.	PVP was bipedicular. For both PVP and BKP, PMMA was injected under x-ray visualisation using a mobile C-arm x-	Barium-opacified PMMA (Zimmer) mixed with an antibiotic (gentamicin);	NR	PVP 4.91±0.65 ml BKP 5.56±0.62 ml

Study	Pre-procedural imaging method used to determine fracture characteristics	Background and experience of operator performing vertebroplasty or kyphoplasty	Anaesthetic	Approach for vertebroplasty or kyphoplasty	Type of cement used	Mean number of vertebrae treated in index procedure	Mean (SD) volume of cement injected (per vertebra)
				ray.	viscosity not specified		
Rousing 2009[139]	Plain x-ray for all patients; MRI (stir weighted) or bone scan (spect) for those with >1 fracture (fractures accepted as new if they showed oedema on MRI or increased bone turn-over on bone scan)	Orthopaedic surgeons specialising in spine surgery	Most patients were mildly conscious sedated; all were prepared for GA in case of complications.	Unipedicular or bipedicular. PMMA injected under continuous biplane fluoroscopy, and injection terminated in case of extravertebral cement leakage.	PMMA: manufacturer and viscosity not specified	Not clear. Probably 1.2	NR
VERTOS[153]	X-ray and MRI for all patients	Not specified	Local anaesthesia	Bipedicular. PMMA injected under continuous fluoroscopy. CT scan with multiplanar reconstruction of the treated levels, performed immediately after vertebroplasty to assess the cement deposition and to identify possible extra cement leakage or other local	PMMA (Osteopal V, Biomet Merck). Viscosity not specified	1.6	3.2 ml (median 3.0, range 1.0-5.0 ml)

Study	Pre-procedural imaging method used to determine fracture characteristics	Background and experience of operator performing vertebroplasty or kyphoplasty	Anaesthetic	Approach for vertebroplasty or kyphoplasty	Type of cement used	Mean number of vertebrae treated in index procedure	Mean (SD) volume of cement injected (per vertebra)
				complications that might not have been noted under fluoroscopy.			
VERTOS II)[17]	X-ray and MRI for all patients	Experienced radiologists[149]	Local anaesthesia	Bipedicular using a single or biplane angiography system under fluoroscopic guidance. Cement was injected under continuous fluoroscopic monitoring to identify local cement leakage or migration into the venous system towards the lungs. (Immediately after the procedure, a CT scan of the treated vertebral bodies was done with 2mm slices to identify cement leakage outside the vertebral body or other possible local complications.[172])	PMMA (Osteo-Firm, COOK Medical). Viscosity not specified	1.3	4.1±1.5 ml (range 1-9 ml)

Table 101: Clinical efficacy outcomes reported

Study	Pain/analgesic use	Back-specific functional status/ mobility	Disability	Vertebral body height/ angular deformity	Progression of treated fracture	Incidence of new fractures	HRQoL	All-cause mortality	Complications	Other adverse events
Blasco 2012[146]	Yes (pain on VAS scale, analgesic use)	No	No	No	No	Yes	Yes (QUALEFFO-41)	Yes	Yes (cement leaks)	No
Buchbinder 2009[101]	Yes (pain on 0-10 scale, opioid use)	Yes (modified RDQ)	No	No	No	Yes	Yes (QUALEFFO, AQoL, EQ-5D)	No	Yes (cement leaks)	Yes
Farrokhi [147]	Yes (pain on 0-10 scale)	Yes (Oswestry LBP scale)	Yes (whether ambulatory on day 1)	Yes (vertebral body height, kyphotic wedge angle)	Yes	Yes	Yes	Yes	Yes (cement leaks)	Yes
FREE[151]	Yes (pain on 0-10 scale, analgesic use)	Yes (RDQ)	Yes (days of restricted activity, use of walking aids, back braces etc)	Yes (kyphotic angle)	Yes	Yes	Yes (SF-36 PCS, EQ-5D)	Yes	Yes (cement leaks)	Yes
INVEST[102]	Yes (pain on 0-	Yes (modified	Yes (SOF-	No	No	No	Yes (SF-36	No	Yes	Yes

Study	Pain/analgesic use	Back-specific functional status/ mobility	Disability	Vertebral body height/ angular deformity	Progression of treated fracture	Incidence of new fractures	HRQoL	All-cause mortality	Complications	Other adverse events
	10 scale, pain frequency index, pain bothersomeness scale, opioid use)	RDQ)	ADL)				PCS, EQ-5D)			
Liu 2010[148]	Yes (pain on VAS scale)	No	No	Yes (vertebral body height, kyphotic wedge angle)	No	Yes (adjacent fractures)	No	No	No	No
Rousing 2009[139]	Yes (pain on VAS scale)	No	Yes (tandem test, timed 'Up & Go' test, repeated chair test)	No	No	Yes	Yes (SF-36 PCS & MCS, DPQ, EQ-5D, Barthel Index, MMSE)	Yes	Yes (cement leaks)	Yes
VERTOS[153]	Yes (pain on 0-10 scale, analgesic use)	Yes (RDQ)	No	No	No	Yes	Yes (QUALEFFO)	Not specifically. Implicitly	Yes	Yes

Study	Pain/analgesic use	Back-specific functional status/ mobility	Disability	Vertebral body height/ angular deformity	Progression of treated fracture	Incidence of new fractures	HRQoL	All-cause mortality	Complications	Other adverse events
								none		
VERTOS II)[17]	Yes (pain on VAS scale, analgesic use)	Yes (RDQ)	No	Yes	Yes (height loss during follow-up of treated fractures[165])	Yes	Yes (EQ-5D, QUALEFFO)	Yes	Yes (cement leaks)	Yes

Table 102: Time points at which clinical efficacy outcomes reported

Study	12-24 hours	3 days	1 week	2 weeks	1 month	2 months	3 months	6 months	12 months	24 months	36 months
Blasco 2012[146]				✓		✓		✓	✓		
Buchbinder 2009[101]			✓		✓		✓	✓			
Farrokhi [147]			✓			✓		✓	✓	✓	✓
FREE[151]					✓		✓*	✓*	✓	✓	
INVEST[102]		✓		✓	✓						
Liu 2010[148]		✓						✓			
Rousing 2009[139]	✓				✓**		✓		✓		
VERTOS[153]	✓			✓							
VERTOS II)[17]	✓		✓		✓		✓	✓	✓		

* very few outcomes

** I month after hospital discharge

Table 103: Reported outcomes compared with outcomes specified in the protocols of included studies

Study	Study protocol available	Clinical outcomes specified in study protocol	Outcomes reported
Blasco 2012[146]	Yes[397]	• Quality of life (QUALEFFO-41 at baseline, 2 weeks, and 2, 6 and 12 months) • Pain (VAS at baseline, 2 weeks, and 2, 6 and 12 months)	• Quality of life (QUALEFFO-41 at baseline, 2 weeks, and 2, 6 and 12 months) • Pain at baseline, 2 weeks, and 2, 6 and 12 months • Analgesic use at baseline, 2 weeks, and 2, 6 and 12 months • Symptomatic vertebral fractures
Buchbinder 2009[101]	Yes[150]	• Pain (overall, at rest, and in bed at night (11-point scale at 1 week, 1, 3, 6, 12, and 24 months) • Quality of life (AQoL, QUALEFFO, EQ-5D at 1 week, 1, 3, 6, 12, and 24 months) • Back pain-related disability (modified RDQ at 1 week, 1, 3, 6, 12, and 24 months) • Timed 'Up and Go' test (at baseline, 12 and 24 months) • Patients' perception of recovery with respect to pain, fatigue, and overall health (7-point ordinal scales at 1 week, 1, 3, 6, 12, and 24 months) • Incidence of new vertebral fractures (radiographs at 12 and 24 months)	• Average pain during 24-hour period; pain at rest, and pain in bed at night (11-point VAS at 1 week, 1, 3, and 6 months) • Quality of life (AQoL, QUALEFFO, EQ-5D at 1 week, 1, 3, and 6 months) • Back pain-related disability (modified RDQ score at 1 week, 1, 1, 3, and 6 months) • Patients' perception of pain at 1 week, 1, 3, and 6 months • Opioid use at 1 week, 1, 3, and 6 months • Adverse events at 1 week, 1, 3, and 6 months
Farrokhi [147]	Yes[197]	• Fast treatment of VCF (measured by radiography at 1 week and 2, 6, 12, 24, and 36 months)	• Average pain during 24-hour period (Huskisson's 10-point scale at 1 week and 2, 6, 12, 24, and 36 months) • Functional quality of life (non-validated Persian

Study	Study protocol available	Clinical outcomes specified in study protocol	Outcomes reported
			translation of Oswestry LBP disability scale at 1 week and 2, 6, 12, 24, and 36 months) • Vertebral body height (measured radiographically at 2, 6, 12, 24, and 36 months) • Sagittal index (measured radiographically at 2, 6, 12, 24, and 36 months) • Mobility on day 1 after start of intervention • Cement leakage • Adverse events
FREE[151]	Yes[398]	• Quality of life (SF-36 PCS at 1 month, EQ-5D and SF-36 at 1, 3, 6, 12 and 24 months) • Function (RDQ and objective functionality tests - reaching, "get up and go" - at 1, 3, 6, 12 and 24 months) • Pain (11-point scale at 5-10 days - post enrollment for the control group and post kyphoplasty for the kyphoplasty group) • Changes in spinal deformity (measured radiographically at baseline, 3, 12 and 24 months) • Maintenance of vertebral body height in Kyphoplasty treated subjects only (lateral spine x-rays at baseline and at 3, 12, and 24 month visits) • Patient satisfaction (1, 3, 6, 12, 24 months) • Outcome (nursing home, back to status prior to fracture, at 1, 3, 6, 12, 24 months), and hospital days, disabilities, etc. at 1, 3, 6, 12, 24 months	• Quality of life (SF-36 PCS and EQ-5D at baseline, 1, 3, 6, 12 and 24 months) • Function (RDQ score at baseline, 1, 3, 6, 12 and 24 months) • Non-pharmacological therapies at baseline, 1 and 12 months • Pain (11-point scale and analgesic use at baseline, 1, 12 and 24 months) • Changes in spinal deformity (postoperatively and at 24 months) • Patient satisfaction at 24 months • Days of restricted activity at 1, 12, and 24 months • Incident fractures at 12 and 24 months • Procedural safety and other adverse events at 12 and 24 months

Study	Study protocol available	Clinical outcomes specified in study protocol	Outcomes reported
		• Rate of incident fractures (frequency, timing and location, at 3, 12 and 24 months) • Procedural safety (peri-operative clinical events)	
INVEST[102]	Yes[152]	• Back pain-related disability (modified RDQ) • Pain (11-point scale, modified Deyo-Patrick Pain Frequency and Bothersomeness Scale) • Analgesic use • Quality of life (SF-36, EQ-5D) • Functional status (SOF-ADL, OPAQ body image domain) • Adjacent fractures (radiograph at 12 months) • Implant-related inflammation (patients receiving vertebroplasty)	• Back pain-related disability (modified RDQ) • Average pain during 24-hour period (11-point scale, modified Deyo-Patrick Pain Frequency and Bothersomeness Scale) • Opioid use • Quality of life (SF-36 PCS and MCS, EQ-5D) • Functional status (SOF-ADL)
Liu 2010[148]	No	-	• Pain on a 10-point scale 3 days and 6 months • Postoperative vertebral body height • Postoperative kyphotic wedge angle • Adjacent fractures
Rousing 2009[139]	No	-	• Pain on a 10 cm VAS at 12-24 hours, 3 and 12 months • Quality of life (SF-36 PCS and MCS at 3 and 12 months; also EQ-5D, Barthel Index, and MMSE in subgroup only) • Effect of pain on daily life (Dallas Pain Questionnaire) • Function (objective functionality tests – tandem

Study	Study protocol available	Clinical outcomes specified in study protocol	Outcomes reported
			test, timed "Up & Go" test, repeated chair test – at 3 and 12 months, in subgroup only) • Incident fractures at 3 and 12 months • Intraoperative cement leakage
VERTOS[153]	No	-	• Back pain recorded on an 11-point scale 1 day and 2 weeks after vertebroplasty or initiation of optimal pain medication • Analgesic use score 1 day and 2 weeks after vertebroplasty or initiation of optimal pain medication • Quality of life (QUALEFFO completed 2 weeks after vertebroplasty or initiation of optimal pain medication • Back pain-related disability (RDQ completed 2 weeks after vertebroplasty or initiation of optimal pain medication)
VERTOS II)[17]	Yes	• Pain (11-point scale at baseline, 1 day, and 1, 3, 6 and 12 months; analgesic use over first month) • Quality of life (QUALEFFO and EQ-5D at baseline, 1 day, and 1, 3, 6 and 12 months) • Back pain-related disability (RDQ) • Secondary fractures (x ray at 1, 3 and 12 months)	• Pain (11-point scale at 1 day, 1 week, and 1, 3, 6 and 12 months; analgesic use at 1 day, 1 week and 1 month) • Quality of life (QUALEFFO and EQ-5D) • Back pain-related disability (RDQ) • Secondary fractures (x ray at 1, 3 and 12 months) • Vertebral body height loss 'during follow-up'[165]

Appendix 9: Clinical efficacy data

Table 104: Mean (SD) AQoL scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: short- and medium-term outcomes (data from Buchbinder et al[101])

Time point	PVP	Control	Difference between groups (95% CI) (positive values favour intervention)	P value
Baseline	0.33 (0.25)	0.27 (0.26)		
1 week	NR	NR		
Change from baseline at 1 week	**0.0 (0.2)**	**0.0 (0.2)**	**0.0 (-0.1 to 0.1)**	**NR**
1 month	NR	NR		
Change from baseline at 1 month	**0.0 (0.2)**	**0.1 (0.3)**	**0.0 (-0.1 to 0.1)**	**NR**
3 months	NR	NR		
Change from baseline at 1 month	**0.0 (0.2)**	**0.1 (0.3)**	**0.0 (-0.1 to 0.1)**	**NR**
6 months	NR	NR		
Change from baseline at 1 month	**0.0 (0.3)**	**0.1 (0.3)**	**0.0 (-0.1 to 0.2)**	**NR**

Table 105: Mean (SD) EQ-5D scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: short- to medium-term outcomes

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
Buchbinder 2009[101]	Baseline	0.30±0.32		0.28±0.33		NR
	1 week	NR		NR		
	Change from baseline at 1 week	**0.1 (0.3)**		**0.1 (0.3)**	**0.0 (-0.1 to 0.2)**	**NR**
	1 month	NR		NR		
	Change from baseline at 1 month	**0.1 (0.3)**		**0.1 (0.3)**	**0.0 (-0.1 to 0.1)**	**NR**
	3 months	NR		NR		
	Change from baseline at 3 months	**0.2 (0.3)**		**0.2 (0.4)**	**0.0 (-0.1 to 0.2)**	**NR**
	6 months	NR		NR		
	Change from baseline at 6 months	**0.2 (0.4)**		**0.2 (0.4)**	**0.0 (-0.1 to 0.2)**	**NR**
FREE[34,151]	Baseline		0.17 (0.37)	0.19 (0.36)		
	1 month		0.59 (0.32)	0.40 (0.33)		
	Change from baseline at 1 month		***0.42***	***0.21***	**0.18 (0.08 to 0.28)‡**	**0.0003**
	3 months		0.62 (0.29)	0.53 (0.33)		
	Change from baseline at 3 months		***0.45 (0.37 to 0.53)***	***0.34 (0.28 to 0.42)***	***0.11 (0.00 to 0.22)***	
	6 months		0.63 (0.31)	0.53 (0.32)		
	Change from baseline at 6 months		***0.46 (0.38 to 0.54)***	***0.34 (0.26 to 0.42)***	***0.12 (0.01 to 0.23)***	
INVEST[102]	Baseline	0.57 (0.18)		0.54 (0.23)		

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
	1 month	0.70 (0.18)		0.64 (0.20)		
	Change from baseline at 1 month	***0.13***		***0.10***	**0.05 (-0.01 to 0.11)**	**0.13**
Rousing 2009[139]	Baseline	(n=17): 0.356 (95% CI 0.196-0.516)		(n=16): 0.083 (95% CI -0.151 to 0.317)		0.05
	3 months	N=15 0.731 (0.653 to 0.809)		N=17 0.543 (0.387 to 0.699)		0.04
	Change from baseline at 3 months	***+0.375 (0.33 to 0.42)***		***+0.460 (0.42 to 0.50)***	***-0.085 (-0.15 to -0.02)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

Table 106: Mean EQ-5D scores (SD) before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: long-term outcomes

Study	Time point	PVP	BKP	Control	Difference between groups (95% CI) (positive values favour intervention)	P value
FREE[34,151,196]	Baseline		0.17 (0.37)	0.19 (0.36)		
	12 months		0.64 (0.29)	0.54 (0.33))		
	Change from baseline at 12 months		***0.47***	***0.35***	**0.12 (0.01 to 0.22)‡**	**0.0252**
	24 months		0.63 (0.29)	0.56 (0.32)		
	Change from baseline at 24 months (mean over 24 months)		***0.46***	***0.37***	**0.12 (0.06 to 0.18)‡**	**0.0002**
Rousing 2009[139,185]	Baseline	N=17: 0.356 (95% CI 0.196-0.516)		N=16: 0.083 (95% CI -0.151 to 0.317)		0.05
	12 months	N=14 0.675 (0.576 to 0.775)		N=18 0.571 (0.448 to 0.694)		0.19
	Change from baseline at 12 months	***+0.319 (0.27 to 0.37)***		***+0.488 (0.45 to 0.52)***	***-0.169 (-0.23 to -0.11)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

Table 107: Change from baseline in mean (SD) EQ-5D scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures

Study	PVP	Control	Adjusted mean between-group difference (95% CI) (positive values favour intervention)	P value
Buchbinder 2009[101]	0.1 (0.3)	0.1 (0.3)	0.0 (-0.1 to 0.1)	NR
INVEST[102]	*0.13*	*0.10*	0.05 (-0.01 to 0.11)	0.13
Pooled data[109]	0.12 (0.19)	0.11 (0.23)	0.03 (-0.02 to 0.08)	NR
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers				

Table 108: Mean (SD) QUALEFFO total scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: short-term outcomes

Study	Time point	PVP	Control	Mean difference between groups (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco pers comm.)	Baseline	65.19 (SE 2.23)	59.17 (SE 2.17)		
	2 weeks	61.16 (SE 2.42)	58.03 (SE 2.29)		
	Change from baseline at 2 weeks	***-4.03 (-10.45 to +2.42)***	***-1.14 (-7.32 to +5.04)***	***-2.89 (-11.74 to +5.96)***	
Buchbinder 2009[101,399]	Baseline	56.9 (13.4)	59.6 (17.1)		
	1 week	NR	NR		
	Change from baseline at 1 week	**-0.5 (7.4)**	**3.6 (9.2)**	**-4.0 (-7.8 to -0.2)†**	
VERTOS[153]	Baseline	60 (range 37-86)	67 (range 38-86)		0.1
	2 weeks	53 (range 28-79)	67 (range 40-88)	-14 (-24.7 to -3.4)	NR
	Change from baseline at 2 weeks	**-6.8**	**-0.7**	**-6.1 (-10.7 to -1.6)**	**NR**
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

† adjusted for stratification and baseline variables

Table 109: Mean (SD) QUALEFFO total scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: medium-term outcomes

Study	Time point	PVP	Control	Mean difference between groups (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco pers comm.)	Baseline	65.19 (SE 2.23)	59.17 (SE 2.17)		
	2 months	57.80 (SE 2.39)	55.65 (SE 2.28)		
	Change from baseline at 2 months	***-7.39 (-13.80 to -0.98)***	***-3.52 (-9.69 to +2.65)***	***-3.87 (-12.62 to +4.88)***	
	6 months	54.13 (SE 2.30)	51.93 (SE 2.25)		
	Change from baseline at 6 months	***-11.06 (-17.34 to -4.78)***	***-7.24 (-13.67 to -1.11)***	***-3.82 (-12.42 to +4.78)***	
Buchbinder 2009[101,399]	Baseline	56.9 (13.4)	59.6 (17.1)		
	1 month	NR	NR		
	Change from baseline at 1 month	**2.8 (9.3)**	**2.4 (12.3)**	**0.9 (-4.2 to 6.0)†**	
	3 months	NR	NR		
	Change from baseline at 3 months	**6.0 (9.6)**	**6.1 (13.7)**	**0.7 (-4.4 to 5.7)†**	
	6 months	NR	NR		
	Change from baseline at 6 months	**6.4 (13.4)**	**6.1 (13.4)**	**0.6 (-5.1 to 6.2)†**	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

† adjusted

Table 110: Mean (SD) QUALEFFO total scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: long-term outcomes

Study	Time point	PVP	Control	Mean difference between groups (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco pers comm.)	Baseline	65.19 (SE 2.23)	59.17 (SE 2.17)		
	12 months	54.38 (SE 2.38)	52.01 (SE 2.32)		
	Change from baseline at 21 months	***-10.81 (-17.20 to -4.42)***	***-7.16 (-13.89 to -0.93)***	***-3.65 (-12.28 to +4.98)***	
VERTOS II)[17]	Baseline	58.7 (13.5)	54.7 (14.4)		>0.05
	1 year	NR	NR	NR	
	Change from baseline at 1 year	NR	NR	NR	<0.0001†
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

† adjusted for baseline differences

Table 111: Mean (SD) SF-36 utility scores before and after percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: medium and long-term outcomes: data from the FREE study[34]

Time point	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
Baseline	[redacted]	[redacted]	[redacted]	
1 month	[redacted]	[redacted]	[redacted]	
Change from baseline at 1 month	[redacted]	[redacted]	[redacted]	
3 months	[redacted]	[redacted]	[redacted]	
Change from baseline at 3 months	[redacted]	[redacted]	[redacted]	
6 months	[redacted]	[redacted]	[redacted]	
Change from baseline at 6 months	[redacted]	[redacted]	[redacted]	
12 months	[redacted]	[redacted]	[redacted]	
Change from baseline at 12 months	[redacted]	[redacted]	[redacted]	
24 months	[redacted]	[redacted]	[redacted]	
Change from baseline at 24 months	[redacted]	[redacted]	[redacted]	

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

Table 112: Mean (SD) SF-36 PCS scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: medium-term outcomes

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
FREE[34,151,196]	Baseline	N/A	[redacted]	[redacted]		
	1 month	N/A	[redacted]	[redacted]		
	Change from baseline at 1 month	**N/A**	[redacted]	[redacted]	**5.2 (2.9 to 7.4)‡**	**<0.0001**
	3 months	N/A	[redacted]	[redacted]		
	Change from baseline at 3 months	**N/A**	[redacted]	[redacted]	**4.0 (1.6 to 6.3)‡**	**0.0008**
	6 months	N/A	[redacted]	[redacted]		
	Change from baseline at 6 months	**N/A**	[redacted]	[redacted]	**3.39 (1.13 to 5.64)‡**	**0.003**
INVEST[102]	Baseline	25.3 (7.8)	N/A	25.3 (7.3)		
	1 month	29.7 (9.6)	N/A	28.7 (8.0)		
	Change from baseline at 1 month	***+4.4***	**N/A**	***+3.4***	**1.0 (-1.7 to 3.7)†**	**0.45**
Rousing 2009[139]	Baseline	36.7 (95% CI 30.0-43.4)	N/A	33.4 (95% CI 26.2-40.7)		
	3 months	34.0 (30.1 to 37.9)	N/A	29.3 (24.5 to 34.1)		0.12
	Change from baseline at 3 months	***-2.7 (-4.52 to -0.88)***	**N/A**	***-4.1 (-6.16 to -2.04)***	***+1.4 (-1.38 to +4.18)***	

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

† Adjusted for baseline value and treatment centre

Table 113: Mean (SD) SF-36 PCS scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: long-term outcomes

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
FREE[34,196]	Baseline		[redacted]	[redacted]		
	12 months		[redacted]	[redacted]		
	Change from baseline at 12 months		[redacted]	[redacted]	**1.70 (-0.59 to 3.98)‡**	**0.15**
	24 months		[redacted]	[redacted]		
	Change from baseline at 24 months		[redacted]	[redacted]	**1.68 (-0.63 to 3.99)‡**	**0.26**
Rousing 2009[139,185]	Baseline	36.7 (95% CI 30.0-43.4)		33.4 (95% CI 26.2-40.7)		
	12 months	32.1 (27.8 to 36.3)		30.5 (25.2 to 35.7)		0.63
	Change from baseline at 12 months	***-4.6 (-6.48 to -2.72)***		***-2.9 (-5.00 to -0.80)***	***-1.7 (-4.57 to +1.17)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

Table 114: Mean (SD) SF-36 MCS scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: medium-term outcomes

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
FREE[34]	Baseline		[redacted]	[redacted]		
	1 month		[redacted]	[redacted]		
	Change from baseline at 1 month		[redacted]	[redacted]	[redacted]	
	3 months		[redacted]	[redacted]		
	Change from baseline at 3 months		[redacted]	[redacted]	[redacted]	
	6 months		[redacted]	[redacted]		
	Change from baseline at 6 months		[redacted]	[redacted]	[redacted]	
INVEST[102]	Baseline	44.8 (11.8)		41.5 (14.1)		
	1 month	46.9 (12.0)		45.6 (14.8)		
	Change from baseline at 1 month	***+2.1***		***+4.1***	**1.0 (-3.7 to 4.6)†**	**0.83**
Rousing 2009[139]	Baseline	49.7 (95% CI 43.6-55.8)		49.6 (95% CI 41.9-57.3)		
	3 months	48.9 (43.8 to 54.0)		46.2 (39.2 to 53.2)		0.51
	Change from baseline at 3 months	***-0.8 (-2.62 to +1.02)***		***-3.4 (-5.84 to -0.96)***	***+2.6 (-0.51 to +5.71)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

† adjusted for baseline value and treatment centre

Table 115: Mean (SD) SF-36 MCS scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: long-term outcomes

Study	Time point	PVP	BKP	Control	Between-group mean difference (95% CI) (positive values favour intervention)	P value
FREE[34]	Baseline		[redacted]	[redacted]		
	12 months		[redacted]	[redacted]		
	Change from baseline at 12 months		[redacted]	[redacted]	[redacted]	
	24 months		[redacted]	[redacted]		
	Change from baseline at 24 months		[redacted]	[redacted]	[redacted]	
Rousing 2009[139,185]	Baseline	49.7 (95% CI 43.6-55.8)		49.6 (95% CI 41.9-57.3)		
	12 months	48.7 (42.7 to 54.6)		49.0 (43.9 to 54.1)		0.93
	Change from baseline at 12 months	***-1.0 (-2.99 to +0.99)***		***-0.6 (-2.70 to +1.57)***	***-0.4 (-3.40 to +2.60)***	

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

Table 116: Mean (SD) RDQ scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: short-term outcomes

Study	Time point	PVP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Buchbinder 2009[101]	Baseline	17.3 (2.8)	17.3 (2.9)		
	1 week	NR	NR		
	Change from baseline at 1 week	**-1.8 (5.0)**	**-4.0 (6.8)**	**-2.1 (-5.2 to 0.9)†**	
INVEST[102]	Baseline	16.6 (3.8)	17.5 (4.1)		
	3 days	13.0 (5.2)	12.5 (5.5)		
	Change from baseline at 3 days	***-3.6***	***-5.0***	**-0.9 (-2.7 to 0.8)††**	**0.30**
	2 weeks	12.4 (5.8)	12.3 (5.9)		
	Change from baseline at 2 weeks	***-4.2***	***-5.2***	**-0.6 (-2.4 to 1.2)††**	**0.35**
VERTOS[153]	Baseline	15.7 (range 8-22)	17.8 (range 9-24)		0.2
	2 weeks	13 (range 3-22)	18 (range 9-23)	-5 (-8.4 to -1.2)	
	Change from baseline at 2 weeks	***-2.7***	***+0.2***	***-2.9****	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

† Adjusted for stratification variables and baseline values

†† Adjusted for baseline value and treatment centre

* CI could not be calculated because neither SD nor SE reported

Table 117: Mean (SD) RDQ scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: medium- and long-term outcomes

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour control)	P value
Buchbinder 2009[101]	Baseline	17.3 (2.8)		17.3 (2.9)		
	1 month	NR		NR		
	Change from baseline at 1 month	**-4.4 (6.6)**		**-3.1 (6.8)**	**-1.7 (-5.2 to +1.8)†**	
	3 months	NR		NR		
	Change from baseline at 3 months	**-3.7 (5.4)**		**-5.3 (7.2)**	**-1.5 (-1.7 to +4.8)†**	
	6 months	NR		NR		
	Change from baseline at 6 months	**-4.1 (5.8)**		**-3.7 (5.8)**	**0.0 (-2.9 to +3.0)†**	
FREE[34,151,196]	Baseline		16.79 (4.95)	17.75 (3.96)		
	1 month		11.36 (6.14)	16.32 (4.46)		
	Change from baseline at 1 month		***-5.43***	***-1.43***	**-4.0 (-5.5 to -2.6)‡**	**<0.0001**
	3 months		10.03 (5.55)	13.40 (6.26)		
	Change from baseline at 3 months		***-6.76***	***-4.35***	***-2.41***	
	6 months		9.37 (5.82)	11.92 (6.17)		
	Change from baseline at 6 months		***-7.42***	***-5.83***	***-1.59***	
	12 months		9.61 (6.24)	12.07 (6.12)		
	Change from baseline at 12 months		***-7.18***	***-5.68***	**-2.6 (-4.1 to -1.0)‡**	**0.0012**
	24 months		9.79 (5.77)	10.89 (6.30)		
	Change from		***-7.00***	***-6.86***	**-1.43**	**0.51**

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour control)	P value
	baseline at 24 months					
INVEST[102]	Baseline	16.6 (3.8)		17.5 (4.1)		
	1 month	12.0 (6.3)		13.0 (6.4)		
	Change from baseline at 1 month	**-4.6**		**-4.5**	**0.7 (-1.3 to 2.8) ††**	**0.49**
VERTOS II[17]	Baseline	18.6 (3.6)		17.2 (4.2)		<0.05
	12 months	NR		NR	NR	
	Change from baseline at 12 months	**NR**		**NR**	**NR**	**<0.0001**

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

† Adjusted for stratification variables and baseline values

†† Adjusted for baseline value and treatment centre

Table 118: Change from baseline in mean (SD) RDQ scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures

Study	PVP	Control	Adjusted mean between-group difference (95% CI) (negative values favour intervention)	P value
Buchbinder 2009[101]	-4.4 (6.6)	-3.1 (6.8)	-1.7 (-5.2 to +1.8)	NR
INVEST[102]	-4.6	-4.5	-0.1	NR
Pooled data[109]	-4.1 (5.9)	-3.9 (6.1)	-0.8 (-0.9 to 2.4)	NR

Table 119: Number of patients in the Buchbinder and INVEST studies showing improvement in RDQ scores at one month: data from Staples et al 2011[109]

Outcome	PVP	Control	Relative risk (95% CI)	P value
Improvement in RDQ score of ≥3 units	49/94 (52.1%)	46/89 (51.7%)	1.0 (0.7 to 1.5)	NS
Improvement in RDQ score of ≥30%	41/102 (40.28%)	41/100 (41.0%)	1.0 (0.7 to 1.4)	NS

Table 120: Mean (SD) Oswestry scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	Control	Mean difference, intervention vs control (95% CI) (negative values favour PVP)	P value
Farrokhi 2011[147]	Baseline	52.2 (2.4)	50.4 (2.8)		
	1 week	30.1 (3.0)	44.0 (2.5)	-14.0 (-15.0 to -12.82)	<0.001
	Change from baseline at 1 week	***-22.1 (-23.29 to -20.91)***	***-6.4 (-7.54 to -5.27)***	***-15.7 (-17.35 to -14.05)***	
	2 months	15.0 (2.2)	30.0 (3.1)	-15.0 (-16.76 to -13.24)	<0.019
	Change from baseline at 2 months	***-37.2 (-38.21 to -36.19)***	***-20.4 (-21.66 to -19.14)***	***-16.8 (-18.43 to -15.17)***	
	6 months	10.0 (2.0)	21.0 (2.5)	-11.0 (-12.17 to -7.83)	<0.011
	Change from baseline at 6 months	***-42.2 (-43.17 to -41.23)***	***-29.4 (-30.54 to -28.27)***	***-12.8 (-14.03 to -11.30)***	
	12 months	8.0 (3.2)	20.0 (1.7)	-12.0 (-13.5 to -11.5)	<0.021
	Change from baseline at 12 months	***-44.2 (-45.46 to -42.94)***	***-30.4 (-31.40 to -29.40)***	***-13.8 (-15.41 to -12.19)***	
	24 months	8.0 (2.2)	20.0 (2.0)	-12.0 (-13.32 to -10.68)	<0.041
	Change from baseline at 24 months	***-44.2 (-45.22 to -43.18)***	***-30.4 (-31.45 to -29.35)***	***-13.8 (-15.26 to -12.34)***	
	36 months	8.0 (1.7)	22.0 (1.2)	-14.0 (-14.91 to -13.09)	<0.01
	Change from baseline at 36 months	***-44.2 (-45.12 to -43.28)***	***-28.4 (-29.33 to -27.47)***	***-15.8 (-17.11 to -14.49)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

Table 121: Mean (95% CI) Barthel Index scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures: medium- and long-term outcomes

Study	Time point	PVP	Control	Mean difference between groups (positive scores favour PVP) (95% CI)	P value
Rousing 2009[139,185]	Baseline	17.7 (15.6-19.8)	17.0 (14.2-19.8)		
	3 months	19.6 (19.0-20.3)	18.1 (16.8-19.4)		0.07
	Change from baseline at 3 months	***1.9 (1.26 to 2.54)***	***1.1 (0.31 to 1.89)***	***+0.8 (-0.23 to +1.83)***	
	12 months	19.8 (19.5-20.0)	18.5 (17.6-19.3)		0.02
	Change from baseline at 12 months	***2.1 (1.49 to 2.71)***	***1.5 (0.75 to 2.25)***	***+0.6 (-0.38 to +1.58)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

Table 122: Mean (SD) SOF-ADL scores before and after percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
INVEST[102]	Baseline	10.0 (3.6)	10.3 (2.8)		
	1 month	7.7 (3.7)	8.2 (3.6)		
	Change from baseline at 1 month	*-2.3*	*-2.1*	0.4 (-0.8 to 1.6)††	0.51
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

†† Adjusted for baseline value and treatment centre

Table 123: Mean SF-36 bodily pain subscale score (SD): data from the FREE study[34]

Time point	BKP	Control	Mean difference, intervention vs control (95% CI) (positive values favour BKP)
Baseline	16.91 (13.95)	15.23 (14.44)	
1 month	44.34 (24.29)	29.49 (17.80)	
Change from baseline at 1 month	***+27.43 (+22.80 to +32.06)***	***+14.26 (+10.41 to +18.11)***	***+13.17 (+7.26 to +19.08)***
3 months	54.48 (27.63)	41.55 (22.57)	
Change from baseline at 3 months	***+37.57 (+32.38 to +42.76)***	***+26.32 (+21.63 to +31.01)***	***+11.25(+4.52 to +17.98)***
6 months	55.03 (26.78)	45.85 (24.91)	
Change from baseline at 6 months	***+38.12 (+33.02 to +43.22)***	***+30.59 (+25.49 to +35.69)***	***+7.53 (+0.65 to +14.41)***
12 months	55.03 (26.78)	49.01 (23.78)	
Change from baseline at 12 months	***+38.12 (+32.90 to +43.34)***	***+33.78 (+28.76 to +38.77)***	***+4.34 (-2.52 to +11.20)***
24 months	57.36 (26.61)	48.66 (22.96)	
Change from baseline at 24 months	***+40.45 (+35.19 to +45.71)***	***+33.43 (+28.59 to +38.27)***	***+7.02 (+0.20 to +13.85)***

Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers

Table 124: Change from baseline in mean (SD) overall pain scores following percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: short-term outcomes

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco pers comm.)	Baseline	7.21 (SE 0.33)		6.31 (SE 0.35)		
	2 weeks	5.87 (SE 0.44)		4.79 (SE 0.41)		
	Change from baseline at 2 weeks	***-1.34 (-2.13 to -0.55)***		***-1.52 (-2.28 to -0.76)***	***+0.18 (-0.95 to +1.31)***	
Buchbinder 2009[101]	Baseline	7.4 (2.1)		7.1 (2.3)		
	1 week	NR		NR		
	Change from baseline at 1 week	**-1.5 (2.5)**		**-2.1 (2.8)**	**+0.7 (-0.4 to 1.8)†**	
Farrokhi 2011[147]	Baseline	8.4 (1.6)		7.2 (1.7)		
	1 week	3.3 (1.5)		6.4 (2.1)	-3.1 (-3.72 to -2.28)	<0.001
	Change from baseline at 1 week	***-5.1 (-5.59 to -4.61)***		***-0.8 (-1.39 to -0.21)***	***-4.3 (-5.11 to -3.49)***	
FREE[34,151]	Baseline		6.85 (1.57)	6.80 (1.53)		
	1 week		NR	NR		
	Change from baseline at 1 week		**NR**	**NR**	**-2.2 (-1.6 to -2.8)‡**	**<0.0001**
INVEST[102]	Baseline	6.9 (2.0)		7.2 (1.8)		
	3 days	4.2 (2.8)		3.9 (2.9)		
	Change from baseline at 3 days	***-2.7***		***-3.3***	**+0.4 (-0.5 to 1.5)††**	**0.37**
	2 weeks	4.3 (2.9)		4.5 (2.8)		
	Change from baseline at 1 week	***-2.6***		***-2.7***	**+0.1 (-0.8 to 1.1)††**	**0.77**
Liu 2010[148]	Baseline	7.9 (0.7)	8.0 (0.8)			
	3 days	2.3 (0.5)	2.6 (0.6)			NS

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	Change from baseline at 3 days	***-5.6 (-5.78 to -5.42)***	***-5.4 (-5.60 to -5.20)***		**PVP vs BKP** ***-0.2 (-0.43 to +0.03)***	
Rousing 2009[139]	Baseline	7.5 (95% CI 6.6-8.4)		8.8 (95% CI 8.2-9.3)		0.02
	12-24 hours	2.0 (0.9 to 3.2)		NR		
	Change from baseline at 12-24 hours	***-5.5***		**NR**	**NR**	
VERTOS[153]	Baseline	7.1 (range 5-9)		7.6 (range 5-10)		0.3
	12-24 hours	4.7 (range 1-8)		7.1 (range 5-10)	-2.4 (-3.7 to -1.0)	
	Change from baseline at 12-24 hours	***-2.4***		***-0.5***	***-1.9****	
	2 weeks	4.9 (range 0-10)		6.4 (range 3-9)	-1.5 (-3.2 to -0.2)	
	Change from baseline at 2 weeks	**-2.1**		**-1.1**	**-1.0 (-0.5 to -2.5)**	
VERTOS II[17]	Baseline	7.8 (1.5)		7.5 (1.6)		
	12-24 hours	3.7 (2.4)		6.7 (2.1)		<0.0001
	Change from baseline at 12-24 hours	***-4.1 (-4.52 to -3.70)***		***-0.8 (-1.19 to -0.41)***	***-3.3 (-3.94 to -2.66)***	
	1 week	3.5 (2.5)		5.6 (2.5)		<0.0001
	Change from baseline at 1 week	***-4.3 (-4.74 to -3.86)***		***-1.9 (-2.35 to -1.45)***	***-2.4 (-3.11 to -1.70)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

† Adjusted for stratification variables and baseline values
‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation
†† Adjusted for study-group assignment, baseline value of the outcome measure, and study centre
* CI could not be calculated because neither SD nor SE reported

Table 125: Change from baseline in mean (SD) overall pain scores following percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: medium-term outcomes

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco [146], Blasco pers comm.)	Baseline	7.21 (SE 0.33)		6.31 (SE 0.35)		
	2 months	4.13 (SE 0.41)		4.72 (SE 0.41)		
	Change from baseline at 2 months	**-3.07 (SE 0.45)**		**-1.59 (SE 0.42)**	***-1.48 (-2.94 to -0.02)***	**0.0172**
	6 months	4.72 (SE 0.36)		4.30 (SE 0.38)		
	Change from baseline at 6 months	***-2.49 (-3.45 to -1.53)***		***-2.01 (-3.02 to -1.00)***	***-0.48 (-1.84 to +0.88)***	
Buchbinder 2009[101]	Baseline	7.4 (2.1)		7.1 (2.3)		
	1 month	NR		NR		
	Change from baseline at 1 month	**-2.3 (2.6)**		**-1.7 (3.3)**	**-0.5 (-1.7 to +0.8)†**	
	3 months					
	Change from baseline at 3 months	**-2.6 (2.9)**		**-1.9 (3.3)**	**-0.6 (-1.8 to +0.7)†**	
	6 months	NR		NR		
	Change from baseline at 6 months	**-2.4 (3.3)**		**-2.1 (3.3)**	**-0.1 (-1.4 to +1.2)**	
Farrokhi 2011[147]	Baseline	8.4 (1.6)		7.2 (1.7)		
	2 months	3.2 (2.2)		6.1 (2.1)	-2.9 (-4.90 to -0.82)	<0.011
	Change from baseline at 2 months	***-5.2 (-6.04 to -4.36)***		***-1.1 (-1.92 to -0.28)***	***-4.1 (-5.28 to -2.92)***	
	6 months	2.2 (2.1)		4.1 (1.5)	-1.9 (-3.25 to -0.55)	<0.021
	Change from baseline at 6 months	***-6.2 (-7.02 to -5.38)***		***-3.1 (-3.79 to -2.41)***	***-3.1 (-4.17 to -2.03)***	
FREE[34,400]	Baseline		6.85 (1.57)	6.80 (1.53)		

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	1 month		3.42 (2.47)	5.33 (2.25)		
	Change from baseline at 1 month		***-3.43***	***-1.47***	**-1.9 (-2.5 to -1.3)‡**	**<0.0001**
	3 months		2.88 (2.57)	4.40 (2.59)		
	Change from baseline at 3 months		***-3.97***	***-2.40***	***-1.57****	
	6 months		2.67 (2.38)	4.24 (2.45)		
	Change from baseline at 6 months		***-4.17***	***-2.56***	***-1.61****	
INVEST[102]	Baseline	6.9 (2.0)		7.2 (1.8)		
	1 month	3.9 (2.9)		4.6 (3.0)		
	Change from baseline at 1 month	***-3.0***		***-2.6***	**-0.7 (-1.7 to 0.3)††**	**0.19**
Liu 2010[148]	Baseline	7.9 (0.7)	8.0 (0.8)			
	~ 6 months	2.6 (0.6)	2.6 (0.6)			NS
	Change from baseline at ~6 months	***-5.3 (-5.56 to -5.04)***	***-5.4 (-5.68 to – 5.12)***		**PVP vs BKP** ***+0.1 (-0.28 to + 0.48)***	
Rousing 2009[139,185]	Baseline	7.5 (95% CI 6.6-8.4)		8.8 (95% CI 8.2-9.3)		0.02
	1 month**	3.4 (2.2 to 4.8)		6.4 (5.0 to 7.9)		0.00
	Change from baseline at 1 month	***-4.1 (-4.46 to -3.74)***		***-2.4 (-2.78 to -2.02)***	***-1.7 (-2.21 to -1.19)***	
	3 months	1.8 (0.8 to 2.8)		2.6 (1.2 to 4.0)		0.33
	Change from baseline at 3 months	***-5.7 (-5.99 to -5.41)***		***-6.2 (-6.52 to -5.88)***	***-0.5 (-0.05 to -0.95)***	
VERTOS II[17]	Baseline	7.8 (1.5)		7.5 (1.6)		
	1 month	2.5 (2.5)		4.9 (2.6)		
	Change from baseline at 1 month	**-5.2 (-4.72 to -5.88)**		**-2.7 (-1.98 to -3.22)**	**-2.6 (-3.37 to -1.74)**	**<0.0001**

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	3 months	2.5 (2.7)		3.9 (2.8)		0.025
	Change from baseline at 3 months	***-5.3 (-5.93 to -4.67)***		***-3.6 (-4.28 to -2.92)***	***-1.7 (-2.62 to -0.78)***	
	6 months	2.3 (2.7)		3.9 (2.9)		0.014
	Change from baseline at 6 months	***-5.5 (-6.14 to -4.86)***		***-3.6 (-4.32 to -2.88)***	***-1.9 (-2.84 to -0.97)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation
† Adjusted for stratification variables and baseline values
†† Adjusted for study-group assignment, baseline value of the outcome measure, and study centre
* CI not calculable
** High risk of bias as data collected retrospectively

Table 126: Change from baseline in mean (SD) overall pain scores following percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: long-term outcomes

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco, pers comm.)	Baseline	7.21 (SE 0.33)		6.31 (SE 0.35)		
	12 months	4.49 (SE 0.39)		4.32 (SE 0.37)		
	Change from baseline at 12 months	***-2.72 (-3.72 to -1.72)***		***-1.99 (-2.99 to -0.99)***	***-0.73 (-2.10 to +0.64)***	
Farrokhi 2011[147]	Baseline	8.4 (1.6)		7.2 (1.7)		
	12 months	2.2 (2.1)		4.1 (1.8)	-1.9 (-2.90 to 0.90)	<0.11
	Change from baseline at 12 months	***-6.2 (-7.03 to -5.37)***		***-3.1 (-3.86 to -2.34)***	***-3.1 (-4.23 to -1.97)***	
	24 months	2.2 (2.1)		3.7 (2.0)	-0.5 (-1.39 to 0.50)	<0.37
	Change from baseline at 24 months	***-6.2 (-7.03 to -5.37)***		***-3.5 (-4.31 to -2.69)***	***-2.7 (-3.86 to -1.54)***	
	36 months	2.8 (2.0)		3.7 (2.5)	-1.5 (-9.85 to 6.85)	<0.81
	Change from baseline at 36 months	***-5.6 (-6.41 to -4.79)***		***-3.5 (-4.44 to -2.56)***	***-2.1 (-3.36 to -0.87)***	
FREE[34,151,196]	Baseline		6.85 (1.57)	6.80 (1.53)		
	12 months		2.70 (2.53)	3.53 (2.41)		
	Change from baseline at 12 months		**NR**	**NR**	**-0.9 (-1.5 to -0.3)‡**	**0.0034**
	24 months		2.61 (2.55)	3.45 (2.49)		

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	Change from baseline at 24 months		**NR**	**NR**	**-0.80 (-1.39 to -0.20)‡**	**0.009**
Rousing 2009[139,185]	Baseline	7.5 (95% CI 6.6-8.4)		8.8 (95% CI 8.2-9.3)		0.02
	12 months	2.0 (1.1 to 3.0)		2.9 (1.6 to 4.1)		0.29
	Change from baseline at 12 months	***-5.5 (-5.79 to -5.21)***		***-5.9 (-6.20 to -5.60)***	***+0.4 (-0.03 to +0.83)***	
VERTOS II[17]	Baseline	7.8 (1.5)		7.5 (1.6)		
	12 months	2.2 (2.7)		3.8 (2.8)		0.014
	Change from baseline at 12 months	**-5.7 (-4.98 to -6.22)**		**-3.7 (-3.05 to -4.35)**	**-2.0 (-2.80 to -1.13)**	**<0.0001**
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

‡ Adjusted for sex, aetiology, current treatment with corticosteroids, and any bisphosphonate treatment within 12 months before randomisation

Table 127: Change from baseline in mean (SD) overall pain scores at one month following percutaneous vertebroplasty for the treatment of osteoporotic vertebral fractures

Study	PVP	Control	Adjusted mean between-group difference (95% CI) (negative values favour intervention)	P value
Buchbinder 2009[101]	-2.3 (2.6)	-1.7 (3.3)	-0.5 (-1.7 to +0.8)	NR
INVEST[102]	*-3.0*	*-2.6*	-0.7 (-1.7 to 0.3)	0.19
Pooled data[109]	-2.8 (3.0)	-2.2 (3.2)	-0.6 (-1.4 to 0.2)	NR
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers				

Table 128: Mean (SD) QUALEFFO pain scores before and after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Blasco 2012 (Blasco, pers comm.)	Baseline	4.03 (SE 0.12)	3.68 (SE 0.12)		
	2 weeks	3.44 (SE 01.4)	3.40 (SE 0.13)		
	Change from baseline at 2 weeks	***-0.59 (-0.95 to -0.23)***	***-0.28 (-0.63 to +0.07)***	***-0.31 (-0.81 to +0.19)***	
	2 months	3.20 (SE 0.14)	3.18 (SE 0.13)		
	Change from baseline at 2 months	***-0.83 (-1.19 to -0.47)***	***-0.50 (-0.85 to -0.15)***	***-0.33 (-0.82 to +0.16)***	
	6 months	3.22 (SE 0.13)	3.12 (SE 0.13)		
	Change from baseline at 6 months	***-0.81 (-1.16 to -0.46)***	***-0.56 (-0.91 to -0.21)***	***-0.25 (-0.73 to +0.23)***	
	12 months	3.05 (SE 0.13)	2.90 (SE 0.13)		
	Change from baseline at 12 months	***-0.98 (-1.33 to -0.63)***	***-0.78 (-1.13 to -0.43)***	***-0.20 (-0.67 to +0.27)***	
Buchbinder 2009[399]	Baseline	72.2 (17.3)	72.1 (16.5)		
	1 week				
	Change from baseline at 1 week	**7.8 (20.5)**	**16.1 (23.1)**	**-8.5 (-18.2 to 1.1)**	
	1 month	NR	NR		
	Change from baseline at 1 month	**14.8 (21.2)**	**19.3 (27.7)**	**-4.0 (-15.1 to 7.1)**	
	3 months	NR	NR		
	Change from baseline at 3 months	**18.1 (21.1)**	**21.1 (30.6)**	**-2.7 (-14.5 to 9.1)**	

Study	Time point	PVP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	6 months	NR	NR		
	Change from baseline at 6 months	**20.4 (25.0)**	**20.7 (25.0)**	**-0.5 (-11.2 to 10.2)**	
VERTOS[153]	Baseline	19	21	-2 (-3.6 to 0.4)	
	2 weeks	14	20	-6 (-8.5 to -2.5)	
	Change from baseline at 2 weeks	***-5***	***-1***	***-4****	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

* CI not calculable

Table 129: Frequency and bothersomeness of pain at one month, compared with baseline; data from the INVEST study

Outcome measure	Time point	PVP	Control	Adjusted treatment effect (95% CI)	P value
Pain Frequency Index	Baseline	3.0±0.8	3.1±0.8		
	1 month	2.1±1.2	2.3±1.1	0.2 (-0.2 to 0.6)	0.33
Pain Bothersomeness Index	Baseline	2.9±0.7	3.1±0.8		
	1 month	1.9±1.1	2.1±1.1	0.2 (-0.2 to 0.6)	0.33

Table 130: Perceived pain: data from Buchbinder et al[101]

Perceived pain	PVP	Control	Relative risk for "better" compared with "no change" or worse"(95% CI)
1 week	**N=37**	**N=37**	
Better	6 (16%)	13 (35%)	0.5 (0.2 to 1.1)
No change	26 (70%)	23 (62%)	
Worse	5 (14%)	1 (3%)	
1 month	**N=35**	**N=38**	
Better	12 (34%)	9 (24%)	1.5 (0.7 to 3.0)
No change	21 (60%)	20 (53%)	
Worse	2 (6%)	9 (24%)	
3 months	**N=36**	**N=37**	
Better	14 (39%)	12 (32%)	1.2 (0.6 to 2.2)
No change	19 (53%)	18 (49%)	
Worse	3 (8%)	7 (19%)	
6 months	**N=35**	**N=36**	
Better	16 (46%)	15 (42%)	1.1 (0.6 to 1.9)
No change	12 (34%)	16 (44%)	
Worse	7 (20%)	5 (14%)	

Table 131: Analgesic use: data from Blasco et al[146]

Time point	No treatment		Minor analgesics		Minor opiate derivatives		Major opiate derivatives	
	PVP	Control	PVP	Control	PVP	Control	PVP	Control
Baseline	9 (14%)	12 (20%)	8 (12%)	17 (28%)	19 (30%)	17 (28%)	28 (44%)	14 (23%)
2 weeks	13 (23%)	12 (21%)	10 (18%)	10 (17%)	13 (23%)	19 (33%)	20 (36%)	17 (29%)
2 months	15 (29%)	16 (29%)	7 (13%)	7 (13%)	14 (27%)	16 (29%)	16 (31%)	17 (30%)
6 months	17 (35%)	13 (25%)	6 (12%)	8 (15%)	8 (16%)	14 (27%)	18 (37%)	17 (33%)
12 months	14 (34%)	17 (40%)	5 (12%)	8 (19%)	7 (17%)	10 (24%)	15 (37%)	7 (17%)

Table 132: Analgesic use: data from VERTOS[153]

	PVP (n=18)	Control (n=16)	Mean difference, intervention vs control (95% CI) (negative values favour intervention)
Pain medication at baseline			
None	2 (11%)	1 (6%)	
Paracetamol	4 (22%)	7 (44%)	
NSAIDs	6 (33%)	3 (19%)	
Opiate derivative	6 (33%)	5 (31%)	
Mean analgesic use score (range)			
Baseline	1.9 (0-3)	1.7 (0-3)	
1 day	1.1 (0-3)	2.5 (1-3)	
Change from baseline at 1 day	**-0.8**	**+0.8**	**-1.6 (-2.3 to -0.8)**
2 weeks	1.2 (0-3)	2.6 (2-3)	
Change from baseline at 2 weeks	**-0.7**	**+0.9**	**-1.5 (-2.3 to -0.8)**

Table 133: Analgesic use at baseline: data from VERTOS II[17]

Analgesic	PVP (n=95)	Control (n=92)
None	5 *(5%)*	7 *(8%)*
Non-opiate drugs	40 *(42%)*	43 *(47%)*
Weak opiate derivatives	31 *(33%)*	22 *(24%)*
Strong opiate derivatives	19 *(20%)*	20 *(22%)*
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers		

Table 134: Mean (SD) vertebral body height (cm) of index fracture after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (positive values favour intervention)	P value
Blasco 2012[146]	Baseline	NR		NR		
	12 months	NR		NR		
	Change from baseline at 12 months	**-0.27 (0.15)**		**-0.13 (0.17)**	***-0.14****	**NS**
Farrokhi 2011[147]	Baseline	2.8 (1.5)		2.5 (1.3)		
	1 week	3.2 (1.1)		2.0 (1.0)	1.2 (1.73 to 0.67)	<0.011
	Change from baseline at 1 week	***+0.4 (-0.18 to +0.98)***		***-0.5 (-1.00 to -0.004)***	***+0.9 (+0.14 to +1.66)***	
	6 months	3.2 (1.1)		1.9 (1.4)	1.3 (2.05 to 0.55)	<0.027
	Change from baseline at 6 months	***+0.4 (-0.18 to +0.98)***		***-0.6 (-1.30 to +0.10)***	***+1.0 (+0.09 to +1.91)***	
	12 months	3.2 (1.5)		2.0 (1.2)	1.2 (2.03 to 0.37)	<0.001
	Change from baseline at 12 months	***+0.4 (-0.27 to +1.07)***		***-0.5 (-1.04 to +0.04)***	***+0.9 (+0.04 to +1.76)***	
	24 months	3.0 (1.5)		2.1 (1.2)	0.9 (1.75 to 0.05)	<0.04
	Change from baseline at 24 months	***+0.2 (-0.47 to +0.87)***		***-0.4 (-0.94 to +0.14)***	***+0.6 (-0.26 to +1.46)***	
	36 months	3.0 (1.2)		2.0 (1.0)	2.0 (1.50 to 0.44)	<0.01
	Change from baseline at 36 months	***+0.2 (-0.41 to +0.81)***		***-0.5 (-1.003 to +0.003)***	***+0.7 (-0.09 to +1.49)***	
Liu 2010[148]	Baseline	1.01 (0.22)	1.13 (0.34)			
	"Postoperative"	1.32 (0.26)	2.04 (0.41)			<0.001

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (positive values favour intervention)	P value
	Change from baseline	***+0.31 (+0.22 to +0.40)***	***+0.91 (+0.76 to +1.06)***		***Favours BKP -0.60 (-0.78 to -0.43)***	
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

* Confidence interval not calculable because baseline data not available

Table 135: Mean (SD) angular deformity of index fracture after percutaneous vertebroplasty or percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
Farrokhi 2011[147]	Baseline	20.0 (5.5)		21.0 (4.2)		
	1 week	10.0 (2.5)		22.0 (2.2)	-12.0 (-12.96 to -11.04)	<0.027
	Change from baseline at 1 week	***-10.0 (-11.87 to -8.13)***		***+1.0 (-0.43 to +2.43)***	***-11.0 (-13.63 to -8.63)***	
	6 months	10.1 (2.6)		23.0 (2.1)	-13.0 (-13.73 to -11.37)	<0.031
	Change from baseline at 6 months	***-9.9 (-11.79 to -8.01)***		***+2.0 (+0.58 to +3.42)***	***-11.9 (-14.27 to -9.54)***	
	12 months	10.0 (1.0)		23.0 (2.0)	-13.0 (-13.47 to -12.53)	<0.001
	Change from baseline at 12 months	***-10.0 (-11.73 to -8.27)***		***+2.0 (+0.58 to +3.42)***	***-12.0 (-14.24 to -9.76)***	
	24 months	9.0 (1.0)		23.0 (2.3)	-14.0 (-14.53 to -13.57)	<0.001
	Change from baseline at 24months	***-11.0 (-12.73 to -9.27)***		***+2.0 (+0.54 to +3.46)***	***-13.0 (-15.26 to -10.74)***	
	36 months	8.9 (1.0)		23.0 (2.0)	-14.0 (-14.96 to -13.04)	<0.011
	Change from baseline at 36 months	***-11.1 (-12.84 to -9.37)***		***+2.0 (+0.58 to +3.42)***	***-13.1 (-15.34 to -10.86)***	
FREE[203]	Baseline		NR	NR		
	Postoperatively		NR	NR		

Study	Time point	PVP	BKP	Control	Mean difference, intervention vs control (95% CI) (negative values favour intervention)	P value
	Change from baseline to "postoperatively"		**-3.3 (-2.4 to -4.2)**	**NR**		
	24 months		NR	NR		
	Change from baseline to 24 months		**-3.1**	**-0.8**	***-2.3****	**0.03**
Liu 2010[148]	Baseline	15.5 (4.2)	17.0 (7.3)			
	"Postoperative"	12.2 (3.6)	9.0 (5.7)			<0.001
	Change from baseline	**-3.3** ***(-5.63 to -0.97)***	**-8.0 (-** ***9.87 to -6.23)***		***BKP vs PVP -4.7 (7.69 to -1.71)***	**<0.001**
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers						

* Confidence interval not calculable because baseline data not available

Table 136: Number of treated vertebrae displaying height loss during follow-up

Height loss during follow-up	PVP (n=136 vertebrae)	Control (n=120 vertebrae)	P
None (0-3 mm)	118	74	<0.001
Moderate (4-7 mm)	7	28	
Severe (≥8 mm)	4	11	

Table 137: All-cause mortality

Study	Percutaneous vertebroplasty	Balloon kyphoplasty	Control	RR (95% CI)	P value
Length of follow-up 2 weeks					
VERTOS[153]	0/18 (0%)	N/A	0/16 (0%)	Not calculable	
Length of follow-up 3 months					
INVEST[102]	0/68 (0%)	N/A	0/63 (0%)	Not calculable	
Length of follow-up 6 months					
Buchbinder 2009[101]	2/38 (5.3%) (chest infection, oesophageal cancer)	N/A	1/40 (2.5%) (acute MI)	*2.11 (0.20 to 22.28)*	*0.54*
Length of follow-up 12 months					
Blasco 2012[146]	3/64 *(4.7%)* (cause not reported)	N/A	6/61 *(9.8%)* (cause not reported)	*0.48 (0.12 to 1.82)*	*0.28*
Rousing 2009[139]	1/25 (4.0%) (cause not reported)	N/A	1/24 (4.2%) (cause not reported)	*0.96 (0.06 to 14.50)*	*0.98*
VERTOS II[17]	5/101 (5.0%) (cardiac failure 4, old age 1)	N/A	6/101 (6.0%) (old age 2, gastric bleeding 1, respiratory insufficiency 1, sepsis 1, cardiac failure 1)	*0.83 (0.26 to 2.64)*	*0.76*
Length of follow-up 24 months					
FREE[196]	N/A	12/149 (8.1%) (cardiovascular 5, pneumonia 1, cancer 3, other 3)	11/151 (7.3%) (cardiovascular 5, pneumonia 2, cancer 2, other 2)	*1.11 (0.50 to 2.43)*	*0.80*
Length of follow-up 36 months					
Farrokhi 2011[147]	2/40 (5.0%) (MI)	N/A	1/42 (2.4%) (cervical cancer)	*2.10 (0.20 to 22.26)*	*0.54*
Data in normal font were taken directly from the text; data in *italics* were calculated by the reviewers					

Appendix 10: Registry data

US registry data

[redacted]

Table 138: Survival in Medicare patients with osteoporotic VCF (data from Exponent report for Medtronic, Appendix E[174])

Treatment	12 months (95% CI)	24 months (95% CI)	48 months (95% CI)
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]

[redacted]

Table 139: Incidence of selected specified AEs in Medicare patients with osteoporotic VCF (data from Exponent report for Medtronic, Appendix E[174])

	12 months (95% CI)	24 months (95% CI)	48 months (95% CI)
Pneumonia			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
Subsequent hospitalisation			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
MI/cardiac complications			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
Pulmonary/ respiratory complications			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
Pulmonary embolism			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
DVT			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
UTI			

Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
Infection			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]
Subsequent VCF with repair			
Non-operated	[redacted]	[redacted]	[redacted]
PVP	[redacted]	[redacted]	[redacted]
BKP	[redacted]	[redacted]	[redacted]

[redacted]

Table 140: Propensity-matched hazard ratios (95% CI) for selected specified AEs in Medicare patients with osteoporotic VCF (data from Exponent report for Medtronic, Appendix F[174])

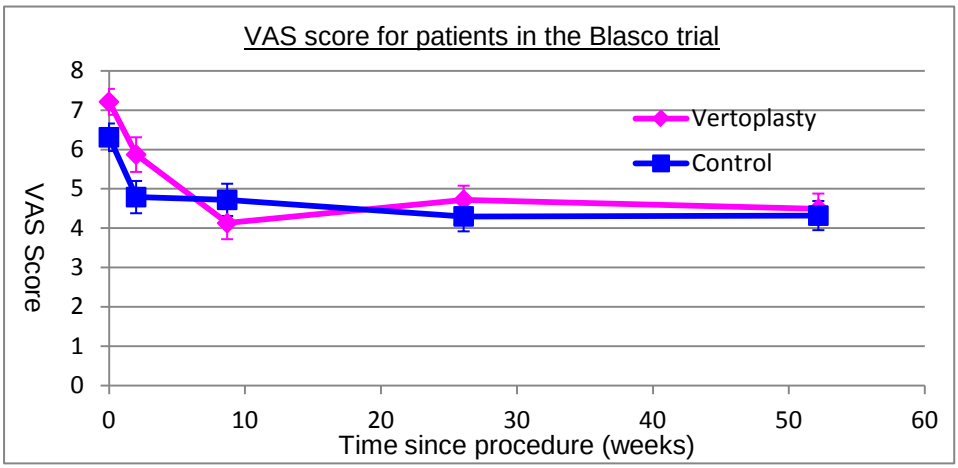
		BKP	VP
Pneumonia			
	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
Subsequent VCF with repair			
	[redacted]	[redacted]	[redacted]
Subsequent hospitalisation			
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
MI/cardiac complications			[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
Pulmonary/ respiratory complications			
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
Pulmonary embolism			
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
DVT			
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
UTI			
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]
Infection			[redacted]
	[redacted]	[redacted]	[redacted]
	[redacted]		[redacted]

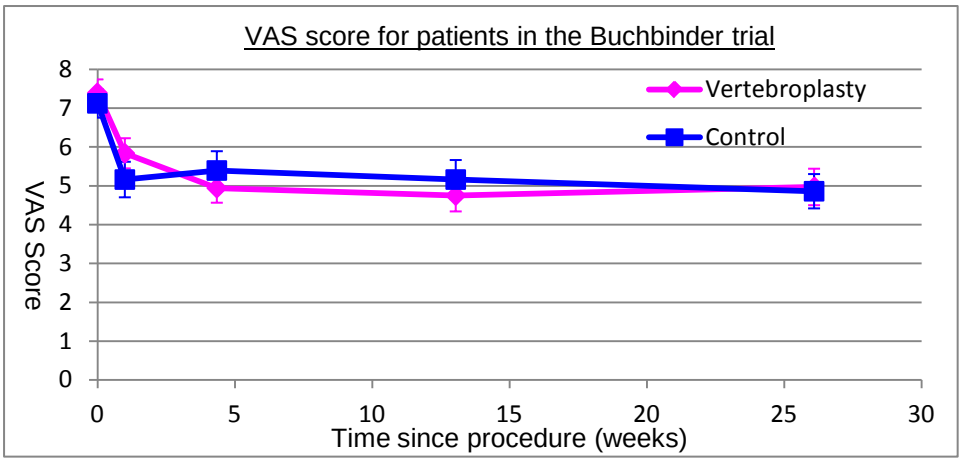
German registry data

[redacted]

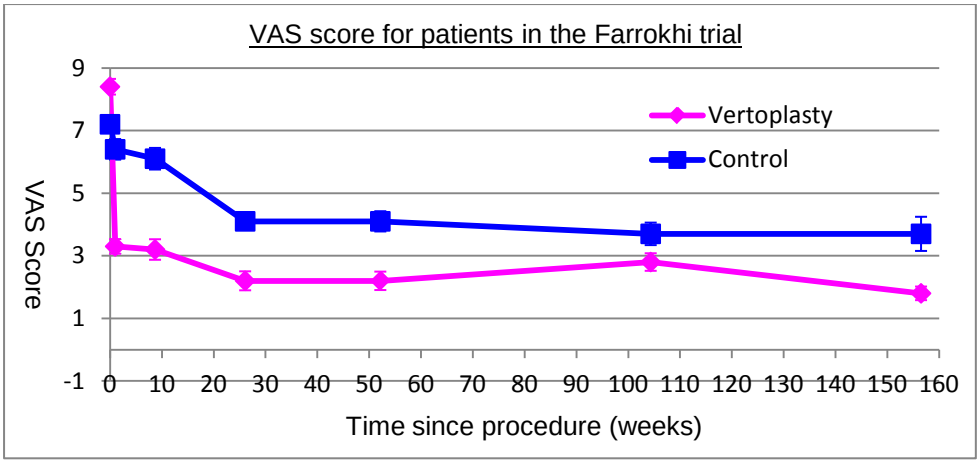
Appendix 11: Longitudinal pain trends

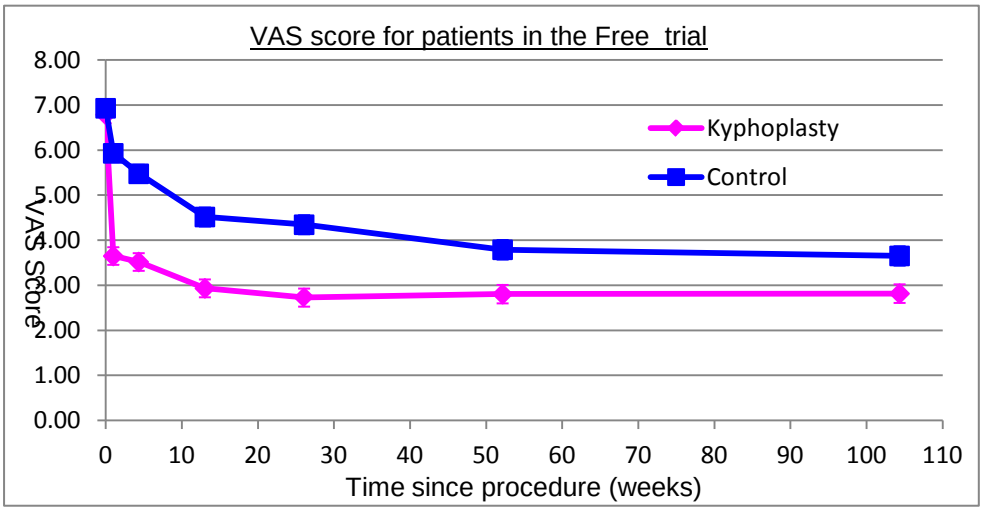
VAS data

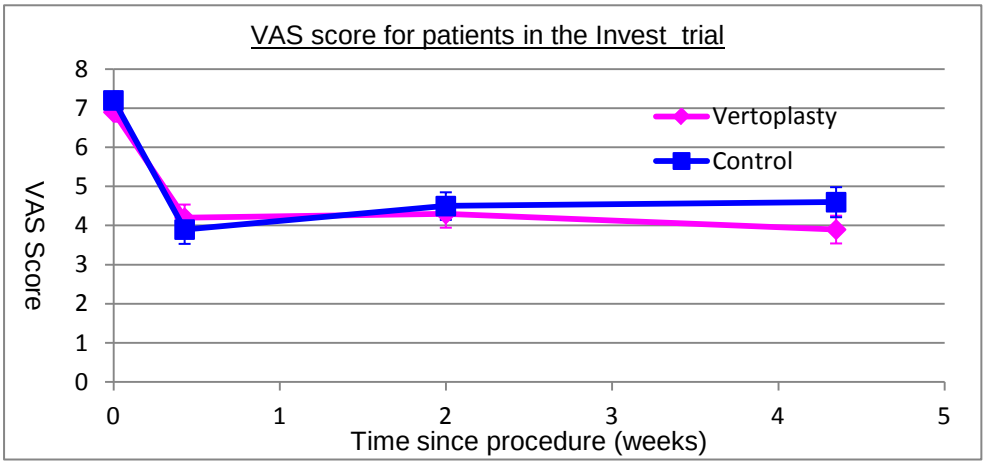


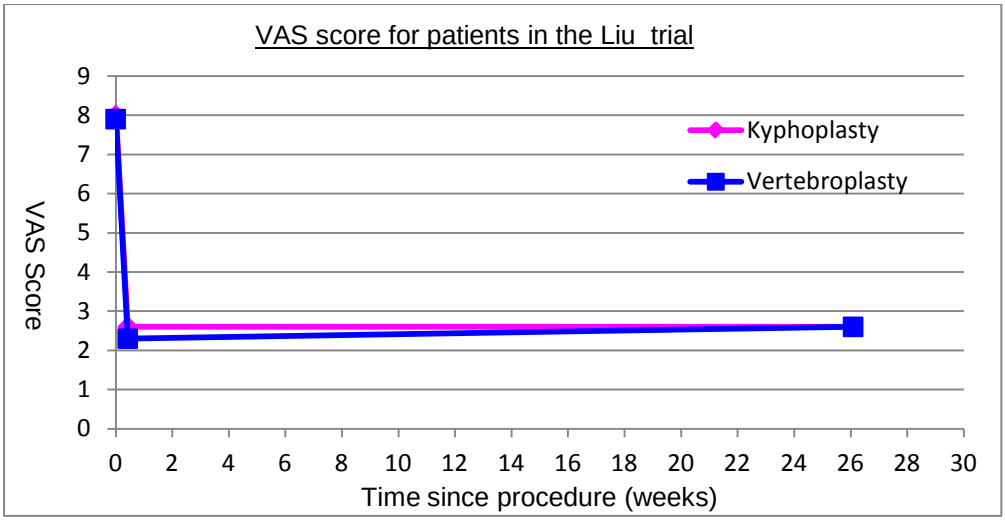


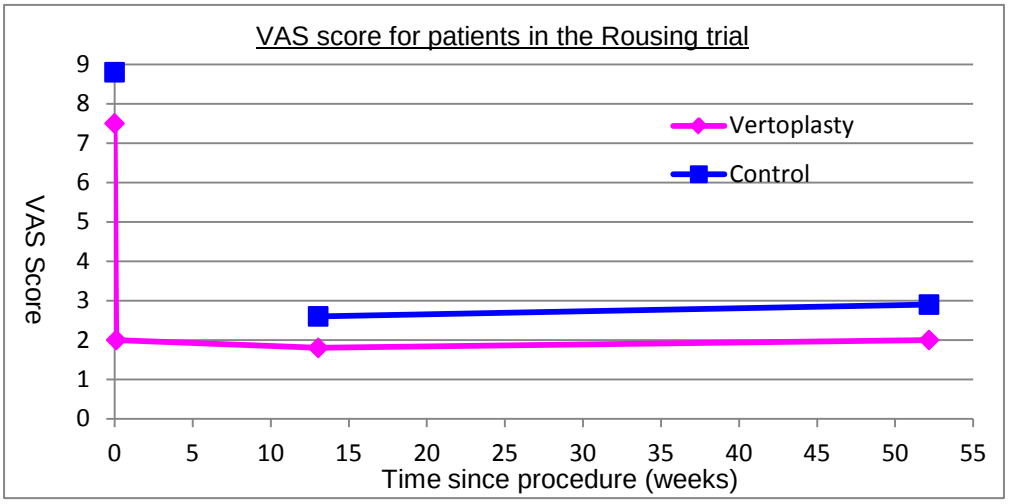
[redacted]

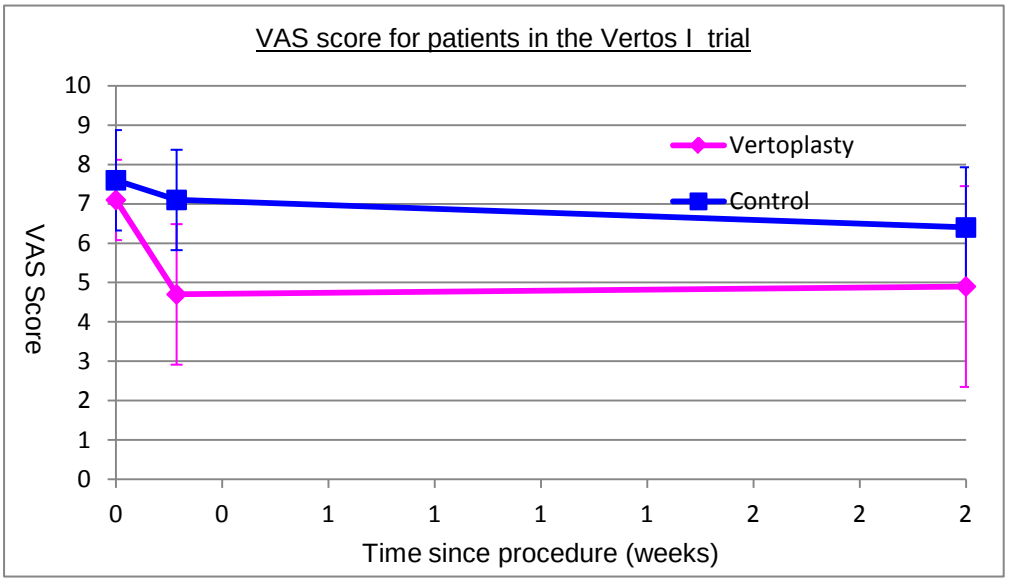

VAS score for patients in the Farrokhi trial
VAS Score
Time since procedure (weeks)
Vertoplasty
Control
9
7
5
3
1
-1
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160

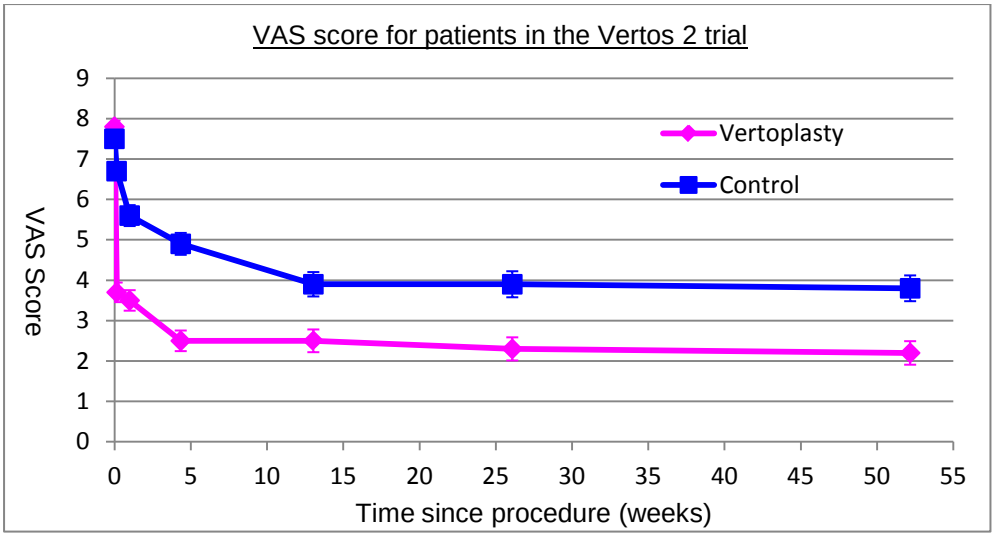

VAS score for patients in the Free trial
VAS Score
Time since procedure (weeks)
Kyphoplasty
Control
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0 10 20 30 40 50 60 70 80 90 100 110


VAS score for patients in the Invest trial
VAS Score
Time since procedure (weeks)
Vertoplasty
Control
8
7
6
5
4
3
2
1
0
0 1 2 3 4 5


VAS score for patients in the Liu trial
VAS Score
Time since procedure (weeks)
Kyphoplasty
Vertebroplasty
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30


VAS score for patients in the Rousing trial
VAS Score
Time since procedure (weeks)
Vertoplasty
Control
0 1 2 3 4 5 6 7 8 9
0 5 10 15 20 25 30 35 40 45 50 55


VAS score for patients in the Vertos I trial
VAS Score
Time since procedure (weeks)
Vertoplasty
Control
0 1 2 3 4 5 6 7 8 9 10
0 0 1 1 1 1 2 2 2


VAS score for patients in the Vertos 2 trial
VAS Score
Time since procedure (weeks)
Vertoplasty
Control
0
1
2
3
4
5
6
7
8
9
0
5
10
15
20
25
30
35
40
45
50
55

APPENDIX 12

REVIEW OF OBSERVATIONAL STUDIES: ESTIMATING MORTALITY DIFFERENCES BETWEEN TREATMENTS FOR VERTEBRAL COMPRESSION FRACTURES.

Dr Monica Hernández: Research Fellow in Econometrics, University of Sheffield

Background: Estimating treatment effects using observational data

The standard problem in treatment evaluation requires the estimation of a causal relationship between the treatment and the outcome. When using observational data, no individual is observed in both the treated and non-treated state and therefore the counterfactual is not observed. Instead a comparison group is generated usually from the same data source as the treated group and used as a control group to estimate the treatment effect. An intuitive and non-technical description of the methods used to evaluate treatment effects when using observational data is found below. A more technical account of the different methods and assumptions used can be found in Cameron and Trivedi (2005).

Different estimators place restrictions on the counterfactuals that can be identified and hence on the treatment effects that can be consistently estimated. There are several treatment effects that might be of interest depending on the evaluation such as the average treatment effect (ATE), the average treatment effect for the treated (ATET) and the average treatment effect for the untreated (ATEU). ATE calculates the expected effect of the treatment if individuals in the population under consideration were randomly allocated to treatment. Depending on the specific evaluation ATE may or may not be the appropriate treatment effect as it is relevant in cases where the treatment is applicable to the entire population represented by the sample data. ATET is relevant when the interest lies on the effect of the treatment for those who are treated and ATEU is relevant when interest lies on the effect of the treatment for those who have not taken the treatment.

Unlike randomised controlled trials observational data are generated in an uncontrolled environment which makes it more difficult to identify the causal relationship of interest. Random assignment to treatment implies that no person is assigned to the treatment on the grounds that the expected benefit is large. In observational data the non-random treatment assignment complicates the estimation of the treatment effect because of selection bias. Selection bias is the difference in the base state between the treated and the control groups and arises when the treatment variable is correlated with the error in the outcome equation. This correlation could be due to incorrectly omitting observable variables that partly determine both the treatment and the outcome (selection on observables) or to the presence of

unobserved factors that partly determine both the treatment and the outcome (selection on unobservables).

The problem of selection on observables can be dealt with by using regression and matching methods. When using regression, the estimated outcome equation needs to include all variables that could be correlated with the error term so there are no omitted observable variables that partly determine the treatment and the outcome. A well specified regression equation identifies the ATE parameter. This method is easy to implement but it usually needs a large set of controls. Furthermore, the functional form employed tends to be quite restrictive and usually assumes that the functions for the treated and non-treated groups conditional on covariates are the same apart from an additional intercept component.

Matching estimators can be used instead of regression to deal with selection on observables. Matching avoids the need of the strong functional form assumptions embedded in a regression model by identifying data from a set of potential comparison individuals (not necessarily from the same population as the treated individuals) with observable characteristics that match those of the treated units up to some specified level of proximity. Matching can be exact or inexact. Exact matching is only possible when there is a small number of discrete covariates and the sample includes a large number of observations for each set of possible covariate values. Inexact matching methods generally use a scalar as the basis for matching individuals. This scalar is obtained as a function of the covariates. The propensity score (Rosebaum and Rubin, 1983) is a popular inexact matching method and it is defined as the conditional probability of receiving treatment given the covariates. To implement matching based on propensity scores, decisions need to be made about whether to match with or without replacement, the number of individuals to use in the control group and what matching method to use. Matching without replacement implies that each observation in the control group is matched to at most one treated observation. Matching with replacement allows multiple matches. If the comparison group is small, the matches may not be very close in terms of the propensity score especially if there is no replacement which will increase the bias of the estimator. When choosing the number of individuals to use in the control group, one needs to take into account the trade off between the bias and the variance of the estimator. The bias is reduced if one uses only the closest match and the variance of the estimator decreases by using more matched controls. However, when using more than one control the bias increases if the extra matches are worse than the first. A compromise solution is to use only matches within a specified radius (calliper matching). In addition, matching is sensitive to the matching method used, the size of the comparison group as well as the amount of overlap between the treated and comparison group. Matching methods are able to identify the ATET parameter.

Matching is more complex than regression and choices about matching methods need to be made and justified in a case by case basis. Matching assumes that the treatment does not indirectly affect untreated individuals (selection into treatment on observables only) and thus requires a good understanding of the process of selection into treatment in each dataset. Matching works best when there exists a rich set of covariates to model this selection, that is, the propensity score. It also assumes that the treated and non-treated groups have comparable observable characteristics (support condition) and needs a large number of potential controls.

Neither regression nor matching methods can identify the parameter of interest if there is selection on unobservables, also referred to as endogeneity of the treatment variable. Endogeneity occurs when a variable which is correlated with the treatment is omitted from the regression. The omission of the variable might be an oversight, or due to the variable not included within the dataset or because the variable cannot be measured. The omitted variable is therefore included in the error term and thus there is an association between the error term and treatment leading to inconsistent estimates of the parameter of interest. In other words, the regression estimate of the parameter measures only the association between the treatment and the outcome rather than the size and direction of the effect. It is often the case that the treatment has a positive effect on the outcome and the unobservable variable is positively correlated with the treatment. This implies that the magnitude of the association obtained using standard regression methods is larger than the true causal effect, that is, the treatment appears more effective than it really is once endogeneity is fully taken into account. In cases like this, Instrumental Variable (IV) methods are needed to identify a parameter related to the ATE, the Local ATE (LATE). The strategy is to choose a variable (the IV variable) that is correlated with the treatment but only correlated with the outcome through its effect on the treatment. Then the variation on this variable can be used to identify the true causal effect due to the treatment only. The estimated LATE is local, because it measures only the effect of the treatment on those that are induced to take the treatment by the change in the IV. The LATE depends on the particular instrument chosen and also on the values of the IV that are used in the estimation. The essential conditions for IV methods to be able to identify the parameter of interest are that the IV is uncorrelated with the error term (also termed exogenous) and that it is correlated with the treatment. If the IV is not strongly correlated with the treatment variable, then the instrument is a weak instrument and the model is said to be weakly identified. It is extremely important that an instrument is exogenous if the instrument is weak. Even mild endogeneity of the instrument can lead to IV parameter estimates that are more inconsistent than those obtained by methods which do not take into account selection on unobservables. Weak instruments in general can also lead to a loss of precision especially on the coefficients of the endogenous variables.

There is no test possible of the hypothesis that the instrument is uncorrelated with the error term. Therefore, exogeneity of the instrument is a subjective decision and has to be justified on theoretical grounds. This often leads to disagreement on instrument validity.

IV methods are relatively easy to implement but they do involve restrictive functional form and identifying assumptions . In practice, however, good exogenous IVs are difficult to find.

Review of included studies

Summary

There were four studies submitted which analysed data relating to mortality differences between non-operated vertebral fracture patients receiving only optimal pain management (OPM) and operated (OP) patients. Further, the group of OP patients is split into patients who undergo percutaneous vertebroplasty (PVP) or those who undergo balloon kyphoplasty (BKP). All four studies use observational data: two of them use the Medicare national inpatient and outpatient claims database in the US and the other two use data from the AOK Niedersachsen health insurance fund in Germany. A variety of methods are used, including Cox regression, matching methods and IV estimation. The results involved paired comparisons between different groups rather than simultaneous comparisons of the three treatments. Only three of the papers report results for the group of patients with osteoporotic vertebral compression fractures (OVCF); one using the Medicare database and two using the German health insurance fund data. Patients who undergo BKP are found to have lower mortality rates than patients who undergo PVP and both groups of operated patients are found to have lower mortality rates than patients who receive OPM. The extent of the difference as well as the level of significance depends on the dataset and methods used although the difference between the group of OP patients and OPM patients is always significant.

Detailed review

Edidin et al (2011) use the Medicare inpatient and outpatient claims between 1 Jan 2005 and 31 Dec 2008 in the US. They analyse the sample of individuals who are 65 or older who did not have a history of vertebral compression fractures (VCF) defined as a VCF diagnosis in the preceding year. After a few other minor exclusions the final sample size is 858,978 of VCF patients which amounts to 85.3% of the total sample of patients with vertebral fractures. A total of 182,946 (21.3%) patients were operated upon and of these 119,253 (13.9%) had a BKP and 63,693 (7.4%) had a PVP. Thus, the number receiving BKP is almost twice the number receiving PVP in this dataset.

Methods: The methods employed include Cox regression (including subgroup analysis for different comorbidities, and patients who survived at least a year following their VCF) and IV methods (two step procedure) with a follow up limited to 3 years due to data requirements for the IV) for differences in mortality. The paper uses a large number of covariates: age; race/ethnithity; patient health status; (general- Charlson comorbidity index groups- and specific – 12 comorbidities that have been

identified previously as possible causes of death associated with VCFs: arterial disease; Chronic obstructive pulmonary disease; cancer; diabetes; hip fracture; hypertensive disease; ischemic heart disease; other heart disease; pneumonia; pulmonary heart disease; stroke; wrist fracture); type of diagnosed fracture (pathologic, traumatic); site of service (outpatient, inpatient); physician specialty (orthopaedic surgeon, neurosurgeon, interventional radiologist, others); socioeconomic status (per capita income for county of residence and Medicare buy-in status); year of diagnosis; and census region (Northeast, Midwest, South and West). The paper also considers four different IVs but only one is finally used: physician preference. Hospital preference, Census region and physician specialty were also given consideration to be used as IV but judged inappropriate.

Results: Using Cox regression a significant difference in survival at 4 years is found between OP and OPM patients as well as between BKP and OPM, PVP and OPM and BKP and PVP patients (see Table 1). Groups with specific comorbidities were analysed separately and the treatment effects were found to be similar. Also the subsample of patients who survived one year after the operation was analysed and although the differences in survival were reduced, they did not completely disappear. Using IV methods and using only the sample of operated patients, they find a relative increase in survival for BKP compared to PVP of 11.82% (at 3 years).

Comments: The analysis uses a sample with both traumatic and osteoporotic fractures and thus assumes that the treatment effect as well as the rest of the estimated parameters are the same for both groups apart from a differential intercept term. In the discussion it is stated that "it remains problematic to attribute a causal relationship between operative treatment and improved patient survival based solely on the results of this study." This is a fair comment and the sensitivity of the results to the assumptions underlying the methods employed should be explored to establish causality and rule out a simple association.

There are some counterintuitive significant associations in the Cox regressions. Arterial disease[1], diabetes, hypertension, ischemic heart disease, stroke and wrist fracture were associated with lower mortality risk which might signal missing variables in the regression. Factors like obesity and smoking were not included due to lack of data, however, as long as they are not correlated with the choice of treatment it should not affect significantly the estimated treatment effect. However, if the missing factors are correlated with the treatment variable, the IV estimator would be a better estimator of the causal effect as long as the IV is exogenous. Unfortunately, the IV used might not be exogenous. Although it is stated that the instruments were "tested" in terms of their correlation with treatment and survival, there is no such test and that is the reason why it is very important to be

[1] Arterial disease is significant in the analysis of OPM vs. OP patients but insignificant in the analysis of BKP vs. PVP.

explicit about the correlation between the IVs considered and the treatment and outcome as this is a subjective decision. It would have been useful to see more information to be able to judge its appropriateness. Some of the instruments that were considered (census region and physician specialty) were included in the regressions which would immediately invalidate them as IVs. A condition for the instrument used, physician preference, to be a good IV is that it should be correlated with the treatment but only correlated with survival through its effect on the selection of treatment. If physician preference of treatment happens because the physician is more likely to get a higher success rate and less complications in the operation, this variable would not be considered a valid IV. This may well be the case in practice in the UK (David Wilson, Consultant Musculoskeletal Interventional Radiologist, Oxford University Hospitals, personal communication).

Exponent (2012) also uses the Medicare inpatient and outpatient claims data in the US but adds an extra year, using data between 1 Jan 2005 and 31 Dec 2009. [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Discussion

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Summing up, it is possible that there is a causal difference in mortality between patients treated using OPM and OP patients given the size of the effect. Appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. It is not possible to say with certainty if there is a difference in mortality between patients undergoing BKP and PVP due to the treatment based on the data presented in the studies included here.

Table 1: Summary of relevant results from included studies.					
Edidin et al (2011) Mortality risk 4 years	Group	Comparison	Cox regression Adjusted HR (95% CI)		IV at 3 years Relative increase In survival
	All	OP vs. OPM	0.63 (0.62-0.64)		
		BKP vs. OPM	0.56 (0.55-0.57)		
		PVP vs. OPM	0.76 (0.75-0.77)		
		BKP vs. PVP	0.77 (0.75-0.78)		
	Survival>1	OP vs. OPM	0.82 (0.81-0.84)		
		BKP vs. OPM	0.76 (0.74-0.77)		
		PVP vs. OPM	0.93 (0.91-0.95)		
		BKP vs. PVP	0.82(0.80-0.85)		
	Operated	BKP vs. PVP			11.82%
Exponent (2012) Mortality risk 5 years	Group	Comparison	Cox regression Adjusted HR (95% CI)		Propensity score Matching and Cox regression HR (95% CI)
	All	OPM vs. OP[(1)]	[redacted]		[redacted]
		OPM vs. BKP	[redacted]		[redacted]
		OPM vs. PVP	[redacted]		[redacted]
		BKP vs. PVP	[redacted]		[redacted]
	OVCF	OPM vs. OP[(1)]	[redacted]		[redacted]
		OPM vs. BKP	[redacted]		[redacted]
		OPM vs. PVP	[redacted]		[redacted]

		BKP vs. PVP	[redacted]		[redacted]
	OVCF Survival>1	OPM vs. OP[(1)]	[redacted]		[redacted]
		OPM vs. BKP	[redacted]		[redacted]
		OPM vs. PVP	[redacted]		[redacted]
		BKP vs. PVP	[redacted]		[redacted]
Lange and Braun (2012a,b) Mortality risk 5 years	Group	Comparison	Cox regression Adjusted HR (95% CI)	Propensity score Matching Difference in survival rates % [p-value]	Propensity score Matching and Cox regression HR (95% CI)
	OVCF	OP vs. OPM	[redacted]		
		BKP vs. PVP	[redacted]	[redacted]	[redacted]
	OVCF Survival>1	OP vs. OPM	[redacted]		
		BKP vs. PVP	[redacted]		

[(1)]Results reported in the appendix but not reported in the main text.

References:

Cameron, AC and Trivedi, PK (2005) Microeconometrics:methods and applications. Cambridge University Press. New York.

Edidin, AA, Ong, KL, Lau, E and Kurtz, SM (2011) Mortality Risk for Operated and Nonoperated Vertebral Fracture Patients in the Medicare Population. Journal of Bone and Mineral Research Vol. 26, No. 7, pp 1617-1626.

Exponent (2012) Mortality and Complication Risks for Operated and Non-Operated Vertebral Fracture Patients in the Medicare Population

Lange, A and Braun, S (January 25, 2012a) Survival analysis using German claims data of patients with Osteoporotic Vertebral Compression Fractures (OVCF). Study Report. Herescon gmbh

Lange, A and Braun, S (February 1, 2012b) Survival analysis using German claims data of patients with Osteoporotic Vertebral Compression Fractures (OVCF). Propensity Score Matching Results (BKP vs. VP). Herescon gmbh

Rosebaum, PR and Rubin, DB (1983) The central role of the propensity score in observational studies for Causal Effects. Biometrika. Vol. 70, No. 1, pp 41-50.

APPENDIX 13

MTC OF MEAN DIFFERENCE IN VAS DURING STABLE PERIOD

Sofia Dias and Tony Ades

SUMMARY

Data consist of the mean VAS at various time points throughout a period of stable pain (after the treatment effect is assumed to have operated), from 4.35 to 156.54 weeks.

A simple analysis of the mean difference in VAS scores over this period was carried out, where the data inputs are the averages of the means reported at all time points and their variances. However, the averaging needs to account for the correlation in the observations at different time points. This correlation was assumed to be 0.87 and constant over time. If other sources of information on the within-study correlation at different time points become available, the calculations can easily be redone.

We describe the method used to impute the within-trial correlations at different time points within the same trial, and for calculating the average and variance of correlated outcomes. The resulting averages and variances are used as data inputs into a standard MTC model in WinBUGS. Results from fixed effects (FE) and random effects (RE) models are described and the FE model is recommended.

There is potential for inconsistency in one loop in this network, but no evidence of inconsistency was found.

DATA

Data on mean VAS score are available from 8 trials, comparing 4 treatments.

Treatments were coded 1 to 4 (Table 1), the data available are described in **TABLE 2** and the network diagram is presented in **FIGURE 1**. OPM was chosen as the overall baseline, or reference treatment.

METHODS

CALCULATING THE MEAN AND VARIANCE OF ALL MEANS IN STABLE PERIOD

For each arm of each study in **TABLE 2**, let y_j be the mean VAS score at time point j and s_j the standard error of the observations at time point j. For a given study, reporting at J time points ($J \geq 1$), we have a vector of observations **Y**, such that,

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{pmatrix} \sim N_J\left(\mathbf{m}, \mathbf{V}\right)$$

where **m** is a vector of unknown means and **V** is the variance-covariance matrix, assumed known. Letting Z represent a linear combination of the elements of **Y**, such that,

$$Z = \frac{y_1 + y_2 + \ldots + y_J}{J} = \mathbf{BY}$$

where

$$\mathbf{B} = \frac{1}{J}\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$$

we have,

$$Var(Z) = \mathbf{BVB'} \qquad (1)$$

For each arm of each study, **V** has in its diagonal the variances of the mean at each time point, s_j^2, and the off-diagonal elements in row *i*, column *j*, will hold $\rho s_j s_i$, where ρ=0.87 and independent of the time lag between observations *i* and *j*.

Repeating this method for all arms of each study, we get $Z = y_{i,k}^*$ the average of the mean VAS score in arm *k* of study *i*, (*i*=1,...,16, *k*=1,2) with variances calculated using equation (1).

The transformed data, on which the MTC will be carried out, are given in **TABLE 3**.

RELATIVE EFFECTS MODEL

The data in **TABLE 3** were used to conduct a MTC, using the model and corresponding WinBUGS code in Dias et al (2011a, Section 3.4).

Briefly, the transformed means are assumed to be normally distributed, so that the likelihood can be written as

$$y_{ik}^* \sim N\left(\theta_{ik}, Var_{ik}\right)$$

The parameter of interest is the mean, θ_{ik}, of this continuous measure which is unconstrained on the real line. The model can be written as

$$\theta_{ik} = \mu_i + \delta_{i,1k} I_{\{k \neq 1\}} \qquad (2)$$

with

$$\delta_{i,1k} \sim N(d_{t_{i1},t_{ik}}, \tau^2)$$

$$I_{\{u\}} = \begin{cases} 1 & \text{if } u \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

where $d_{t_{i1},t_{ik}}$ represents the mean effect of the treatment in arm k in trial i, t_{ik}, compared to the treatment in arm 1 of trial i, t_{i1}, and τ^2 represents the between-trial variability in treatment effects (heterogeneity). Under the exchangeability (consistency) assumption we can write

$$d_{t_{i1},t_{ik}} = d_{1,t_{ik}} - d_{1,t_{i1}}$$

For a FE model we replace equation (2) with

$$\theta_{ik} = \mu_i + d_{t_{i1},t_{ik}} I_{\{k \neq 1\}}$$

Non-informative N(0,100^2) priors are given to the μ's and d's. In a RE model a Uniform(0,10) prior was used for τ.

RESULTS

Model fit statistics for the FE and RE models are given in **TABLE 4**. Although the RE model has a slightly better fit, this is at the expense of more parameters and the DIC does not favour any of the models. We will therefore prefer the FE model, due to its simplicity and easier interpretation. However, a re-examination of the Blasco study (study 1) is recommended as it is showing a relatively poor fit – this appears to be because it is the only study comparing OPM to Vertebroplasty and (marginally) favouring OPM, while all other trials favour Vertebroplasty.

A plot of the effects (mean differences) of all treatments relative to each other is given in **FIGURE 2**. Differences > 0 favour the lowest numbered treatment.

Vertebroplasty and Kyphoplasty (treatments 2 and 3) appear to be the best treatments, although the differences are small (about 1 point) and may not be clinically significant. (For more results see attached WinBUGS files)

CONSISTENCY

There is only one evidence loop in this network (**FIGURE 1**) formed by treatments 1,2,3. Consistency was checked by comparing the treatment effects obtained from separate pairwise meta-analysis for each pair of treatments using the Bucher approach as recommended in Dias et al. (2011b). No evidence of inconsistency was found (Bayesian p-value > 0.8).

REFERENCES

Dias, S., Welton, N. J., Sutton, A. J., & Ades, A. E. 2011a, *NICE DSU Technical Support Document 2: A generalised linear modelling framework for pair-wise and network meta-analysis*, NICE Decision Support Unit, available from http://www.nicedsu.org.uk/.

Dias, S., Welton, N. J., Sutton, A. J., & Ades, A. E. 2011b, *NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials*, NICE Decision Support Unit, available from http://www.nicedsu.org.uk/.

TABLES

TABLE 1 TREATMENT CODES. TREATMENT 1 IS ASSUMED TO BE THE BASELINE REFERENCE TREATMENT TO WHICH ALL OTHERS ARE COMPARED.

Treatment code	Treatment name	Treatment code	Treatment name
1	OPM	3	Kyphoplasty
2	Vertebroplasty	4	OPLA

TABLE 2 DATA AVAILABLE FOR STABLE PERIOD. 'N/A' DENOTES DATA NOT AVAILABLE. TREATMENT CODES ARE GIVEN IN TABLE 1.

ID	Weeks	T1	T2	T1		T2		Trial
				Mean	SE	Mean	SE	
1	8.70	1	2	n/a	n/a	4.13	0.41	Blasco
1	26.09	1	2	4.30	0.38	4.72	0.36	Blasco
1	52.18	1	2	4.32	0.37	4.49	0.39	Blasco
2	4.35	2	4	4.94	0.38	5.40	0.50	Buchbinder
2	13.04	2	4	4.75	0.41	5.16	0.50	Buchbinder
2	26.09	2	4	4.97	0.47	4.86	0.44	Buchbinder
2	52.18	2	4	[redacted]	[redacted]	[redacted]	[redacted]	Buchbinder
2	104.36	2	4	[redacted]	[redacted]	[redacted]	[redacted]	Buchbinder
3	8.70	1	2	n/a	n/a	3.20	0.33	Farrokhi
3	26.09	1	2	4.10	0.26	2.20	0.30	Farrokhi
3	52.18	1	2	4.10	0.32	2.20	0.29	Farrokhi
3	104.36	1	2	3.70	0.36	2.80	0.28	Farrokhi
3	156.54	1	2	3.70	0.55	1.80	0.22	Farrokhi
4	4.35	1	3	n/a	n/a	3.52	0.20	Free
4	13.04	1	3	4.52	0.21	2.93	0.20	Free
4	26.09	1	3	4.35	0.21	2.73	0.20	Free
4	52.18	1	3	3.79	0.22	2.81	0.20	Free
4	104.36	1	3	3.65	0.21	2.82	0.21	Free
5	4.35	2	4	3.90	0.35	4.60	0.38	Invest
6	13.04	1	2	2.60	0.71	1.80	0.51	Rousing
6	52.18	1	2	2.90	0.64	2.00	0.48	Rousing
7	4.35	1	2	n/a	n/a	2.50	0.26	Vertos II
7	13.04	1	2	3.90	0.30	2.50	0.28	Vertos II
7	26.09	1	2	3.90	0.32	2.30	0.29	Vertos II
7	52.18	1	2	3.80	0.32	2.20	0.29	Vertos II
8	26.09	2	3	2.60	0.08	2.60	0.08	Liu

TABLE 3 TRANSFORMED MEANS AND VARIANCES FOR MEAN VAS SCORE FOR INPUT INTO WINBUGS

treatments		Number of arms	Data				RefID
arm 1	arm 2		arm 1		arm 2		
t[,1]	t[,2]	na[]	y[,1]	Var[,1]	y[,2]	Var[,2]	
1	2	2	4.310	0.131	4.447	0.137	Blasco
2	4	2	4.702	0.161	5.114	0.190	Buchbinder
1	2	2	3.900	0.126	2.440	0.072	Farrokhi
1	3	2	4.078	0.041	2.962	0.037	Free
2	4	2	3.900	0.123	4.600	0.144	Invest
1	2	2	2.750	0.426	1.900	0.229	Rousing
1	2	2	3.867	0.090	2.375	0.071	Vertos II
2	3	2	2.600	0.006	2.600	0.006	Liu

TABLE 4 MODEL FIT STATISTICS FOR MTC ANALYSES

	resdev*	pD	DIC	heterogeneity (τ)			
				mean	sd	median	CrI
RE	16.3	13.8	30.1	0.52	0.46	0.42	(0.02,1.64)
FE	18.5	11.0	29.5				

* compare to 16 data points

FIGURES

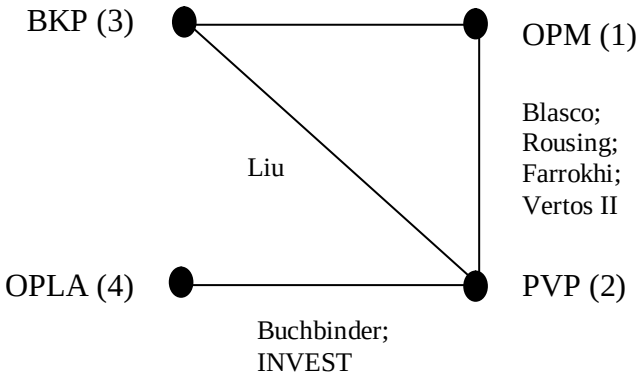


FIGURE 1 TREATMENT NETWORK FOR MEAN VAS SCORE IN STABLE PERIOD.

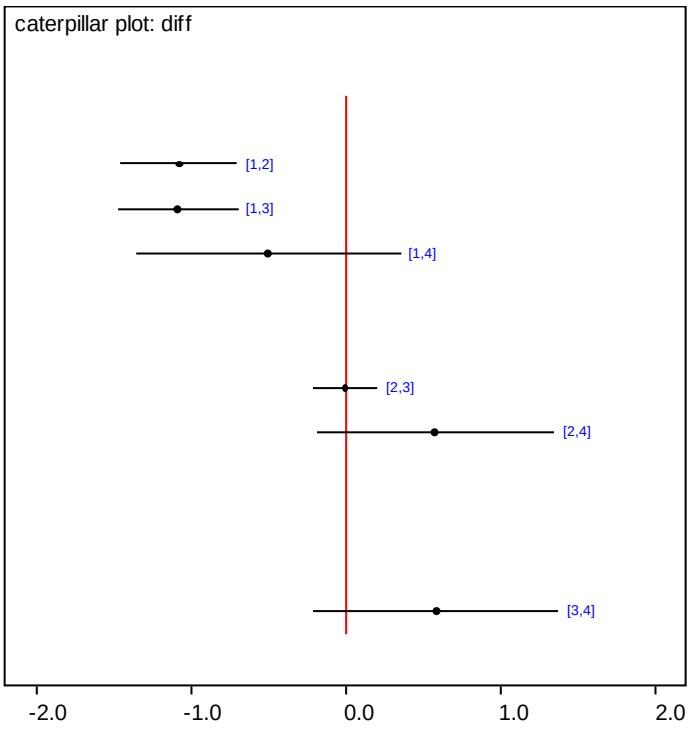


FIGURE 2 PLOT OF MEAN DIFFERENCES OF ALL TREATMENTS RELATIVE TO EACH OTHER. VALUES TO THE RIGHT OF THE HORIZONTAL (RED) LINE FAVOUR THE LOWEST NUMBERED TREATMENT. TREATMENT CODES ARE GIVEN IN TABLE 1.

11. REFERENCES

1. Glaser, D.L., Kaplan, F.S. Osteoporosis. Definition and clinical presentation. *Spine (Phila Pa 1976)* 1997; 22(24 Suppl):12S-16S.

2. Kanis, J.A. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. WHO Study Group. *Osteoporos Int* 1994; 4(6):368-381.

3. Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *American Journal of Medicine* 1993; 94:646-650.

4. Cummings, S.R., Melton, L.J., III. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002; 359:1761-1762.

5. Lindsay, R., Silverman, S.L., Cooper, C., Hanley, D.A., Barton, I., Broy, S.B. et al. Risk of new vertebral fracture in the year following a fracture. *JAMA* 2001; 285:320-323.

6. Holroyd, C., Cooper, C., Dennison, E. Epidemiology of osteoporosis. *Best Practice & Research Clinical Endocrinology & Metabolism* 2008; 22(5):671-678.

7. Genant, H.K., Wu, C.Y., van Kuijk, C., Nevitt, M.C. Vertebral fracture assessment using a semiquantitative technique. *Journal of Bone and Mineral Research* 1993; 8(9):1137-1148.

8. Fink, H.A., Milavetz, D.L., Palermo, L., Nevitt, M.C., Cauley, J.A., Genant, H.K. et al. What proportion of incident radiographic vertebral deformities is clinically diagnosed and vice versa? *Journal of Bone & Mineral Research* 2005; 20(7):1216-1222.

9. Kado, D.M., Browner, W.S., Palermo, L., Nevitt, M.C., Genant, H.K., Cummings, S.R. Vertebral fractures and mortality in older women: a prospective study. Study of Osteoporotic Fractures Research Group. *Arch Intern Med* 1999; 159(11):1215-1220.

10. Ross, P.D. Clinical consequences of vertebral fractures. *American Journal of Medicine* 1997; 103(2A):30S-42S.

11. Lyritis, G.P., Mayasis, B., Tsakalakos, N., Lambropoulos, A., Gazi, S., Karachalios, Th. et al. The natural history of the osteoporotic vertebral fracture. *Clinical Rheumatology* 1989; 8(Suppl 2):66-69.

12. De Smet, A.A., Robinson, R.G., Johnson, B.E., Lukert, B.P. Spinal compression fractures in osteoporotic women: patterns and relationship to hyperkyphosis. *Radiology* 1988; 166:497-500.

13. Melton, L.J. Epidemiology of spinal osteoporosis. *Spine* 1997; 22(24S):2S-11S.

14. Kamath, S., Venkatanarasimha, N., Silver, D.A.T. Percutaneous vertebroplasty. *CPD Journal Radiology Update* 2007; 6(2):82-96.

15. Silverman, S.L. The clinical consequences of vertebral compression fracture. *Bone* 1992; 13(Suppl 2):27-31.

16. Gangi, A., Sabharwal, T., Irani, F.G., Buy, X., Morales, J.P., Adam, A. et al. Quality assurance guidelines for percutaneous vertebroplasty. *Cardiovascular & Interventional Radiology* 2006; 29(2):173-178.

17. Klazen, C.A.H., Lohle, P.N.M., Jansen, F.H., Tielbeek, A.V., Blonk, M.C., Venmans, A. et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. *Lancet* 2010; 376:1085-1092.

18. Burton, A.W., Rhines, L.D., Mendel, E. Vertebroplasty and kyphoplasty: a comprehensive review. *Neurosurgical Focus* 2005; 18(3):e1.

19. Klazen, C.A., Verhaar, H.J., Lampmann, L.E., Juttmann, J.R., Schoemaker, M.C., van Everdingen, K.J. et al. Clinical course of pain in acute osteoporotic vertebral compression fracture. *Journal of Vascular & Interventional Radiology* 2010; 21(9):1405-1409.

20. van Staa, T.P., Dennison, E.M., Leufkens, H.G.M., Cooper, C. Epidemiology of fractures in England and Wales. *Bone* 2001; 29(6):517-522.

21. Aebi, M. Vertebroplasty: about sense and nonsense of uncontrolled "controlled randomized prospective trials". *European Spine Journal* 2009; 18(9):1247-1248.

22. Schlaich, C., Minne, H.W., Bruckner, T., Wagner, G., Gebest, H.J., Grunze, M. et al. Reduced pulmonary function in patients with spinal osteoporotic fractures. *Osteoporosis International* 1998; 8:261-267.

23. Culham, G.E., Jimenez, H.A., King, C.E. Thoracic kyphosis, rib mobility, and lung volumes in normal women and women with osteoporosis. *Spine* 1994; 19:1250-1255.

24. Leech, J.A., Dulberg, C., Kellie, S., Pattee, L., Gay, J. Relationship of lung function to severity of osteoporosis in women. *American Review of Respiratory Disease* 1990; 141(1):68-71.

25. Melton, L.J., III, Kallmes, D.F. Epidemiology of vertebral fractures: implications for vertebral augmentation. [Review] [58 refs]. *Academic Radiology* 2006; 13(5):538-545.

26. Kanis, J.A., Oden, A., Johnell, O., De Laet, C., Jonsson, B. Excess mortality after hospitalisation for vertebral fracture. *Osteoporosis International* 2004; 15:108-112.

27. Cooper, C., Atkinson, E.J., Jacobsen, S.J., O'Fallon, M., Melton, L.J., III. Population-based study of survival after osteoporotic fractures. *American Journal of Epidemiology* 1993; 137:1001-1005.

28. Browner, W.S., Seeley, D.G., Vogt, T.M., Cummings, S.R. Non-trauma mortality in elderly women with low bone mineral density. Study of Osteoporotic Fractures Research Group. *Lancet* 1991; 338:355-358.

29. Cauley, J.A., Thompson, D.E., Ensrud, K.C., Scott, J.C., Black, D. Risk of mortality following clinical fractures. *Osteoporosis International* 2000; 11:556-561.

30. Ström, O., Leonard, C., Marsh, D., Cooper, C. Cost-effectiveness of balloon kyphoplasty in patients with symptomatic vertebral compression fractures in a UK setting. *Osteoporosis International* 2010; 21(9):1599-1608.

31. Ström, O., Borgström, F., Kanis, J.A., Compston, J., Cooper, C., McCloskey, E.V. et al. Osteoporosis: burden, health care provision and opportunities in the EU. *Archives of Osteoporosis* 2011; 6:59-155.

32. Synthes GmbH. Percutaneous stentoplasty (vertebral body stenting) for the treatment of osteoporotic vertebral fractures. Synthes submission for the NICE Multiple Technology Appraisal on osteoporotic vertebral fractures. 2012.

33. Wroth, C., Wiles, A. Key population and vital statistics. Population and vital statistics by area of usual residence in the United Kingdom, 2007. 2009; PPI No 30. Office for National Statistics.

34. Medtronic. Percutaneous vertebroplasty (PVP) and balloon kyphoplasty (BKP) for the treatment of osteoporotic vertebral compression fractures (OVCFs). 2012.

35. Johnson & Johnson. Vertebroplasty and kyphoplasty for the treatment of osteoporotic vertebral fractures. A submission by Johnson & Johnson for the CONFIDENCE SPINAL CEMENT SYSTEM^{TM}. 2012.

36. Singer, B.R., McLauchlan, G.J., Robinson, C.M., Christie, J. Epidemiology of fractures in 15 000 adults. The influence of age and gender. *Journal of Bone & Joint Surgery - British Volume* 1998; 80-B(243):248.

37. Gold, D.T. The clinical impact of vertebral fractures: quality of life in women with osteoporosis. *Bone* 1996; `8(3 Suppl):185S-189S.

38. Lyles, K.W., Gold, D.T., Shipp, K.M., Pieper, C.F., Martinez, S., Mulhausen, P.L. Association of osteoporotic vertebral compression fractures with impaired functional status. *American Journal of Medicine* 1993; 94(6):595-601.

39. Kado, D.M., Browner, W.S., Palermo, L., Nevitt, M.C., Genant, H.K., Cummings, S.R. Vertebral fractures and mortality in older women. A prospective study. *Archives of Internal Medicine* 1999; 159:1215-1220.

40. Ettinger, B., Black, D.M., Palermo, L., Nevitt, M.C., Melnikoff, S., Cummings, S.R. Kyphosis in older women and its relation to back pain, disability and osteopenia: the Study of Osteoporotic Fractures. *Osteoporosis International* 1994; 4:55-60.

41. Ensrud, K.E., Nevitt, M.C., Yunis, C., Cauley, J.A., Seeley, D.G., Fox, K.M. et al. Correlates of impaired function in older women. *Journal of the American Geriatrics Society* 1994; 9(9):1429-1432.

42. Cook, D.J., Guyatt, G.H., Adachi, J.D., Clifton, J., Griffith, L.E., Epstein, R.S. et al. Quality of life issues in women with vertebral fractures due to osteoporosis. *Arthritis & Rheumatism* 1993; 36(6):750-756.

43. Greendale, G.A., Barrett-Connor, E., Ingles, S., Haile, R. Late physical and functional effects of osteoporotic fracture in women: the Rancho Bernardo Study. *Journal of the American Geriatrics Society* 1995; 43(9):955-961.

44. Ryan, P.J. A clinical profile of back pain and disability in patients with spinal osteoporosis. *Bone* 1994; 15(1):27-30.

45. Ettinger, B., Block, J.E., Smith, R., Cummings, S.R., Harris, S.T., Genant, H.K. An examination of the association between vertebral deformities, physical disabilities and psychosocial problems. *Maturitas* 1988; 10(4):283-296.

46. Scane, A.C., Sutcliffe, A.M., Francis, R.M. The sequelae of vertebral crush fractures in men. *Osteoporosis International* 1994; 4:89-92.

47. Dolan, P., Torgerson, D. The cost of treating osteoporotic fractures in the UK female population. *Osteoporosis International* 1998; 8:611-617.

48. Puffer, S., Torgerson, D.J., Sykes, D., Brown, P., Cooper, C. Health care costs of women with symptomatic vertebral fractures. *Bone* 2004; 35(2):383-386.

49. Cooper, C., Atkinson, E.J., O'Fallon, W.M., Melton, L.J., III. Incidence of clinically diagnosed vertebral fractures: a population-based study in Rochester, Minnesota, 1985-1989. *Journal of Bone & Mineral Research* 1992; 7(2):221-227.

50. Ioannidis, G., Papaioannou, A., Hopman, W.M., Akhtar-Danesh, N., Anastassiades, T., Pickard, L. et al. Relation between fractures and mortality: results from the Canadian Multicentre Osteoporosis Study. *CMAJ* 2009; 181(5):265-271.

51. Woolf, A.D., Pfleger, B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003; 81(9):646-656.

52. Kanis, J.A. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* 2002; 359(9321):1929-1936.

53. Marshall, D.A., Johnell, O., Wedel, H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 1996; 312:1254-1259.

54. Prather, H., Hunt, D., Watson, J.O., Gilula, L.A. Conservative care for patients with osteoporotic vertebral compression fractures. *Physical Medicine & Rehabilitation Clinics of North America* 2007; 18(3):577-591.

55. Maksymowych, W.P. Managing acute osteoporotic vertebral fractures with calcitonin. *Canadian Family Physician* 1998; 44:2160-2166.

56. Doidge, J., Merlin, T., Liufu, Z., Tamblyn, D., Jia, L.Y., Hiller, J. Review of interim funded service: vertebroplasty and new review of kyphoplasty. 2011. Commonwealth of Australia.

57. Suzuki, N., Ogikubo, O., Hansson, T. The course of the acute vertebral body fragility fracture: its effect on pain, disability and quality of life during 12 months. *European Spine Journal* 2008; 17:1380-1390.

58. Denaro, L., Longo, U.G., Denaro, V. Vertebroplasty and Kyphoplasty: Reasons for Concern? *Orthopedic Clinics of North America* 2009; 40(4):465-471.

59. Nairn, R.J., Binkhamis, S., Sheikh, A. Current perspectives on percutaneous vertebroplasty: current evidence/controversies, patient selection and assessment, and technique and complications. *Radiology Research and Practice* 2011; 2011:Article ID 175079.

60. Hu, S.S. Internal fixation in the osteoporotic spine. *Spine* 1997; 22(24 Suppl):43S-48S.

61. Jensen, M.E. Percutaneous vertebroplasty: a new therapy for the treatment of painful vertebral body compression fractures. *Applied Radiology* 2000; 29(6):7-12.

62. Frey, M.E. Redo kyphoplasty with vertebroplasty technique: a case report and review of the literature. *Pain Physician* 2009; 12(3):645-649.

63. Novartis. Miacalcin®. *Internet* 2011; Available from http://www.pharma.us.novartis.com/product/pi/pdf/miacalcin_nasal.pdf (accessed Apr. 12 A.D.).

64. Stevenson, M., Davis, S., Kanis, J.A. The hospitalization costs and outpatient costs of fragility fractures. *Women's Health Medicine* 2006; 3(4):149-151.

65. Borgstrom, F., Jonsson, B., Strom, O., Kanis, J.A. An economic evaluation of strontium ranelate in the treatment of osteoporosis in a Swedish setting: based on the results of the SOTI and TROPOS trials. *Osteoporos Int* 2006; 17(12):1781-1793.

66. Stevenson, M., Lloyd Jones, M. The cost effectiveness of an RCT comparing alendronate with Vitamin K1. *Medical Decision Making* 2011; 31:43-52.

67. Francis, R.M., Aspray, T.J., Hide, G., Sutcliffe, A.M., Wilkinson, P. Back pain in osteoporotic vertebral fractures. *Osteoporosis International* 2008; 19(7):895-903.

68. Papaioannou, A., Watts, N.B., Kendler, D.L., Yuen, C.K., Adachi, J.D., Ferko, N. Diagnosis and management of vertebral fractures in elderly adults. [Review] [94 refs]. *American Journal of Medicine* 2002; 113(3):220-228.

69. Rapado, A. General management of vertebral fractures. *Bone* 1996; 18(3 SUPPL.):191-196.

70. Brown, F.L., Jr., Bodison, S., Dixon, J., Davis, W., Nowoslawski, J. Comparison of diflunisal and acetaminophen with codeine in the treatment of initial or recurrent acute low back strain. *Clin Ther* 1986; 9 Suppl C:52-58.

71. Innes, G.D., Croskerry, P., Worthington, J., Beveridge, R., Jones, D. Ketorolac versus acetaminophen-codeine in the emergency department treatment of acute low back pain. *J Emerg Med* 1998; 16(4):549-556.

72. Tramer, M.R., Moore, R.A., Reynolds, D.J., McQuay, H.J. Quantitative estimation of rare adverse events which follow a biological progression: a new model applied to chronic NSAID use. *Pain* 2000; 85(1-2):169-182.

73. Tederko, P., Kiwerski, J., Barcinska-Wierzejska, I. Complex management of osteoporotic vertebral fracture. *Ortop Traumatol Rehabil* 2002; 4(2):157-163.

74. Jellema, P., van Tulder, M.W., van Poppel, M.N., Nachemson, A.L., Bouter, L.M. Lumbar supports for prevention and treatment of low back pain: a systematic review within the framework of the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2001; 26(4):377-386.

75. Longo, U.G., Loppini, M., Denaro, L., Maffulli, N., Denaro, V. Osteoporotic vertebral fractures: current concepts of conservative care. *Br Med Bull* 2012; 102:171-189.

76. Tamayo-Orozco, J., Arzac-Palumbo, P., Peon-Vidales, H., Mota-Bolfeta, R., Fuentes, F. Vertebral fractures associated with osteoporosis: patient management. [Review] [20 refs]. *American Journal of Medicine* 1997; 103(2A):44S-48S.

77. Furlan, A.D., Imamura, M., Dryden, T., Irvin, E. Massage for low-back pain. *Cochrane Database of Systematic Reviews* 2008;(4).

78. French, S.D., Cameron, M., Walker, B.F., Reggars, J.W., Esterman, A.J. Superficial heat or cold for low back pain. *Cochrane Database of Systematic Reviews* 2012;(1).

79. Brewer, L., Williams, D., Moore, A. Current and future treatment options in osteoporosis. *Eur J Clin Pharmacol* 2011; 67(4):321-331.

80. Bolland, M.J., Avenell, A., Baron, J.A., Grey, A., MacLennan, G.S., Gamble, G.D. et al. Effect of calcium supplements on risk of myocardial infarction and cardiovascular events: meta-analysis. *BMJ* 2010; 341:c3691.

81. Torgerson, D.J., Bell-Syer, S.E. Hormone replacement therapy and prevention of vertebral fractures: a meta-analysis of randomised trials. *BMC Musculoskelet Disord* 2001; 2:7.

82. Tavani, A., La, V.C. The adverse effects of hormone replacement therapy. *Drugs Aging* 1999; 14(5):347-357.

83. Wilson, D.J., Owen, S., Corkill, R.A. Facet joint injections as a means of reducing the need for vertebroplasty in insufficiency fractures of the spine. *European Radiology* 2011; 21(8):1772-1778.

84. Oxfordshire Priorities Forum (NHS Oxfordshire). INTERIM Policy Statement 154: Percutaneous Vertebroplasty for Osteoporotic Vertebral Fractures. 2010; available from http://www.oxfordshirepct.nhs.uk/professional-resources/priority-setting/lavender-statements/documents/LS154_Vertebroplasty.pdf (accessed June 2012).

85. National Institute for Clinical Excellence. Interventional Procedure Guidance 12: percutaneous vertebroplasty. 2003.

86. National Institute for Health and Clinical Excellence. Interventional procedure guidance 166: Balloon kyphoplasty for vertebral compression fractures. 2006.

87. Ploeg, W.T., Veldhuizen, A.G., The, B., Sietsma, M.S. Percutaneous vertebroplasty as a treatment for osteoporotic vertebral compression fractures: a systematic review. *European Spine Journal* 2006; 15(12):1749-1758.

88. National Institute for Clinical Excellence Interventional Procedures Programme. Interventional procedure overview of percutaneous vertebroplasty (methyl methacrylate). 2003.

89. Wiles, M.D., Nowicki, R.W.A., Hancock, S.M., Boszczyk, B. Anaesthesia for vertebroplasty and kyphoplasty. *Current Anaesthesia and Critical Care* 2009; 20(1):38-41.

90. Nussbaum, D.A., Gailloud, P., Murphy, K. A review of complications associated with vertebroplasty and kyphoplasty as reported to the Food and Drug Administration medical device related web site.. *Journal of Vascular & Interventional Radiology* 2004; 15(11):1185-1192.

91. Hill, J.M. Vertebroplasty: a new radiographic technique for treating painful spinal compression fractures. *Images* 2001; 20(1):6-8.

92. Goto, K., Hashimoto, M., Takadama, H., Tamura, J., Fujibayashi, S., Kawanabe, K. et al. Mechanical, setting, and biological properties of bone cements containing micron-sized titania particles. *Journal of Materials Science-Materials in Medicine* 2008; 19(3):1009-1016.

93. Wren, A.W., Coughlan, A., Placek, L., Towler, M.R. Gallium containing glass polyalkenoate anti-cancerous bone cements: glass characterization and physical properties. *J Mater Sci Mater Med* 2012; 23(8):1823-1833.

94. Whitehouse, R.W. (i) Patient selection criteria for vertebroplasty or kyphoplasty in painful osteoporotic fracture. *Orthopaedics and Trauma* 2011; 25(2):79-82.

95. Medtronic. Kyphon® balloon kyphoplasty procedure. *Medtronic* 2012; (accessed Feb. 2012).

96. Sandberg, J. Medtronic launches bone cement with hydroxyapatite in the United States. KYPHON® ActivOs™10 bone cement is the latest offering from Medtronic for treatment of vertebral compression fractures. *Internet* 2010; Available from http://www.orthospinenews.com/medtronic-launches-bone-cement-with-hydroxyapatite-in-the-united-states-kyphon%C2%AE-activos%E2%84%A210-bone-cement-is-the-latest-offering-from-medtronic-for-treatment-of-vertebral-compression-fractur (accessed Feb. 2012).

97. Rotter, R., Martin, H., Fuerderer, S., Gabl, M., Roeder, C., Heini, P. et al. Vertebral body stenting: a new method for vertebral augmentation versus kyphoplasty. *European Spine Journal* 2010; 19(6):916-923.

98. Destouet, J.M., Gilula, L.A., Murphy, W.A., Monsees, B. Lumbar facet joint injection: indication, technique, clinical correlation, and preliminary results. *Radiology* 1982; 145(2):321-325.

99. Peh, W. Image-guided facet joint injection. *Biomed Imaging Interv J* 2011; 7(1):e4.

100. Ryan, P.J., Evans, P., Gibson, T., Fogelman, I. Osteoporosis and chronic back pain: a study with single-photon emission computed tomography bone scintigraphy. *J Bone Miner Res* 1992; 7(12):1455-1460.

101. Buchbinder, R., Osborne, R.H., Ebeling, P.R., Wark, J.D., Mitchell, P., Wriedt, C. et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *New England Journal of Medicine* 2009; 361(6):557-568.

102. Kallmes, D.F., Comstock, B.A., Heagerty, P.J., Turner, J.A., Wilson, D.J., Diamond, T.H. et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *New England Journal of Medicine* 2009; 361(3):569-579.

103. FDA. Class II special controls guidance document: polymethylmethacrylate (PMMA) bone cement: guidance for industry and FDA. *Internet* 2011; Available from http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm072795.htm (accessed Nov. 2011).

104. FDA Center for Devices and Radiological Health. Class II special controls guidance document: Polymethylmethacrylate (PMMA) bone cement; guidance for industry and FDA. *Internet* 2002; Available from http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm072837.pdf (accessed Nov. 2011).

105. National Institute for Health and Clinical Excellence. Interventional procedures overview of balloon kyphoplasty for vertebral compression fractures. 2005.

106. Bono, C.M., Heggeness, M., Mick, C., Resnick, D., Watters, W.C. North American Spine Society. Newly released vertebroplasty randomized controlled trials: a tale of two trials. *Spine Journal* 2010; 10(3):238-240.

107. Clark, W.A., Diamond, T.H., McNeil, H.P., Gonski, P.N., Schlaphoff, G.P., Rouse, J.C. Vertebroplasty for painful acute osteoporotic vertebral fractures: Recent Medical Journal of Australia editorial is not relevant to the patient group that we treat with vertebroplasty. *Medical Journal of Australia* 2010; 192(6):334-337.

108. Rad, A.E., Kallmes, D.F. Correlation between preoperative pain duration and percutaneous vertebroplasty outcome. *AJNR Am J Neuroradiol* 2011; 32(10):1842-1845.

109. Staples, M.P., Kallmes, D.F., Comstock, B.A., Jarvik, J.G., Osborne, R.H., Buchbinder, R. Effectiveness of vertebroplasty using individual patient data from two randomised placebo controlled trials: meta-analysis. *BMJ* 2011; 343:d3952.

110. Hansen, G.R., Streltzer, J. The psychology of pain. *Emerg Med Clin North Am* 2005; 23(2):339-348.

111. Bolster, M.B. Consternation and questions about two vertebroplasty trials. *Cleveland Clinic Journal of Medicine* 2010; 77(1):12-16.

112. Gangi, A., Clark, W.A. Have recent vertebroplasty trials changed the indications for vertebroplasty?. [Review]. *Cardiovascular & Interventional Radiology* 2010; 33(4):677-680.

113. Riew, K.D., Yin, Y., Gilula, L., Bridwell, K.H., Lenke, L.G., Lauryssen, C. et al. The effect of nerve-root injections on the need for operative treatment of lumbar radicular pain. A prospective, randomized, controlled, double-blind study. *J Bone Joint Surg Am* 2000; 82-A(11):1589-1593.

114. Noonan, P. Randomized Vertebroplasty Trials: Bad News or Sham News? *American Journal of Neuroradiology* 2009; 30(10):1808-1809.

115. Kaptchuk, T.J. Placebo studies and ritual theory: a comparative analysis of Navajo, acupuncture and biomedical healing. *Philos Trans R Soc Lond B Biol Sci* 2011; 366(1572):1849-1858.

116. Green, S.A. Surgeons and shamans: the placebo value of ritual. *Clin Orthop Relat Res* 2006; 450:249-254.

117. American Academy of Orthopaedic Surgeons. The treatment of symptomatic osteoporotic spinal compression fractures. Guideline and evidence report. 2010. 2011.

118. Ostelo, R.W.J.G., Deyo, R.A., Stratford, P., Waddell, G., Croft, P., Von Korff, M. et al. Interpreting changes scores for pain and functional status in low back pain. *Spine* 2008; 33(1):90-94.

119. Bombardier, C. Outcome assessments in the evaluation of treatment of spinal disorders. Summary and general recommendations. *Spine* 2000; 25(24):3100-3103.

120. Hawthorne, G., Osborne, R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. *Australian and New Zealand Journal of Public Health* 2005; 29(2):136-142.

121. Longo, U.G., Loppini, M., Denaro, L., Maffulli, N., Denaro, V. Rating scales for low back pain. *British Medical Bulletin* 2010; 94:81-144.

122. Walters, S.J., Brazier, J.E. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of Life Research* 2005; 14:1523-1532.

123. Medical Advisory Secretariat. Percutaneous vertebroplasty for treatment of painful osteoporotic vertebral compression fractures: an evidence-based analysis. *Ontario Health Technology Assessment Series* 2010; 10(19):1-45.

124. Ware, J.E.Jr. SF-36 health survey update. *Spine* 2000; 25(24):3130-3139.

125. Copay, A.G., Glassman, S.D., Subach, B.R., Berven, S., Schuler, T.C., Carreon, L.Y. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *The Spine Journal* 2008; 8:968-974.

126. Angst, F., Aeschlimann, A., Stucki, G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis Care & Research* 2001; 45:384-391.

127. Wiebe, S., Matijevic, S., Eliasziw, M., Derry, P.A. Clinically important change in quality of life in epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry* 2002; 73:116-120.

128. Lips, P., Leplège, A. Development and validation of a quality of life questionnaire for patients with vertebral fractures: Qualeffo-41. *Quality of Life Research* 2000; 9:763-766.

129. Oleksik, A., Lips, P., Dawson, A., Minshall, M.E., Shen, W., Cooper, C. et al. Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *Journal of Bone & Mineral Research* 2000; 15(7):1384-1392.

130. van Schoor, N.M., Knol, D.L., Glas, C.A., Ostelo, R.W., Leplège, A., Cooper, C. et al. Development of the Qualeffo-31, an osteoporosis-specific quality-of-life questionnaire. *Osteoporosis International* 2006; 17(4):543-551.

131. Bombardier, C. Outcome assessments in the evaluation of treatment of spinal disorders. *Spine* 2000; 25(24):3097-3099.

132. Trout, A.T., Kallmes, D.F., Gray, L.A., Goodnature, B.A., Everson, S.L., Comstock, B.A. et al. Evaluation of vertebroplasty with a validated outcome measure: the Roland-Morris Disability Questionnaire. *Ajnr: American Journal of Neuroradiology* 2005; 26(10):2652-2657.

133. Furlan, A.D., Pennick, V., Bombardier, C., van Tulder, M. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009; 34(18):1929-1941.

134. Roland, M., Fairbank, J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000; 25(24):3115-3124.

135. Kopec, J.A. Measuring functional outcomes in persons with back pain. A review of back-specific questionnaires. *Spine* 2000; 25(24):3110-3114.

136. Lauridsen, H.H., Hartvigsen, J., Manniche, C., Korsholm, L., Grunnet-Nilsson, N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskeletal Disorders* 2006; 7:82.

137. Patrick, D.L., Deyo, R.A., Atlas, S.J., et al. Assessing health-related quality of life in patients with sciatica. *Spine* 1995; 20:1899-1908.

138. Mahoney, F.I., Barthel, D.W. Functional evaluation: the Barthel Index. *Maryland State Medical Journal* 1965; 14:56-61.

139. Rousing, R., Andersen, M.O., Jespersen, S.M., Thomsen, K., Lauritsen, J. Percutaneous vertebroplasty compared to conservative treatment in patients with painful acute or subacute osteoporotic vertebral fractures. Three-months follow-up in a clinical randomized study. *Spine* 2009; 34(13):1349-1354.

140. Guralnik, J.M., Ferrucci, L., Simonsick, E.M., Salive, M.E., Wallace, R.B. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *New England Journal of Medicine* 1995; 332:556-561.

141. Von Korff, M., Jensen, M.P., Karoly, P. Assessing global pain severity by self-report in clinical and health services research. *Spine* 2000; 25(24):3140-3151.

142. Johnson, C. Measuring pain. Visual analog scale versus numeric pain scale: what is the difference? *Journal of Chiropractic Medicine* 2005; 4(1):43-44.

143. DeLoach, L.J., Higgins, M.S., Caplan, A.B., Stiff, J.L. The Visual Analog Scale in the immediate postoperative period: intrasubject variabiltiy and correlation with a numeric scale. *Anesthesia & Analgesia* 1998; 86:102-106.

144. Bolton, J.E., Wilkinson, R.C. Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients. *Journal of Manipulative & Physiological Therapeutics* 1998; 21(1):1-7.

145. Grotle, M., Brox, J.A., Vøllestad, N.K. Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine* 2004; 29(21):E492-E501.

146. Blasco, J.A., Martinez-Ferrer, A., Macho Fernández, J., San Roman Manzanera, L., Pomés Talló, J., Carrasco Jordan, J.Ll. et al. Effect of vertebroplasty on pain relief, quality of life and the incidence of new vertebral fractures. A 12-month randomised follow-up, controlled trial. *Journal of Bone and Mineral Research* 2012; Accepted manuscript online.

147. Farrokhi, M.R., Alibai, E., Maghami, Z. Randomized controlled trial of percutaneous vertebroplasty versus optimal medical management for the relief of pain and disability in acute osteoporotic vertebral compression fractures. *Journal of Neurosurgery Spine* 2011; 14(5):561-569.

148. Liu, J.T., Liao, W.J., Tan, W.C., Lee, J.K., Liu, C.H., Chen, Y.H. et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. *Osteoporosis International* 2010; 21(2):359-364.

149. Klazen, C.A.H., Verhaar, H.J.J., Lampmann, L.E.H., Juttmann, J.R., Blonk, M.C., Jansen, F.H. et al. VERTOS II: percutaneous vertebroplasty versus conservative therapy in patients with painful osteoporotic vertebral compression fractures; rationale, objectives and design of a multicenter randomized controlled trial. *Trials* 2007; 8(33).

150. Buchbinder, R., Osborne, R.H., Ebeling, P.R., Wark, J.D., Mitchell, P., Wriedt, C.J. et al. Efficacy and safety of vertebroplasty for treatment of painful osteoporotic vertebral fractures: a randomised controlled trial [ACTRN012605000079640]. *BMC Musculoskeletal Disorders* 2008; 9:156.

151. Wardlaw, D., Cummings, S.R., Van Meirhaeghe, J., Bastian, L., Tillman, J.B., Ranstam, J. et al. Efficacy and safety of balloon kyphoplasty compared with non-

surgical care for vertebral compression fracture (FREE): a randomised controlled trial. *Lancet* 2009; 373:1016-1024.

152. Gray, L.A., Jarvik, J.G., Heagerty, P.J., Hollingworth, W., Stout, L., Comstock, B.A. et al. INvestigational Vertebroplasty Efficacy and Safety Trial (INVEST): a randomized controlled trial of percutaneous vertebroplasty. *BMC Musculoskeletal Disorders* 2007; 8:126.

153. Voormolen, M.H.J., Mali, W.P.T.M., Lohle, P.N.M., Fransen, H., Lampmann, L.E.H., van der Graaf, Y. et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteoporotic vertebral compression fractures. The VERTOS study. *American Journal of Neuroradiology* 2007; 28(3):555-560.

154. McKiernan, F., Faciszewski, T., Jensen, R. Reporting height restoration in vertebral compression fractures. *Spine* 2003; 28(22):2517-2521.

155. McKiernan, F.E. Kyphoplasty and vertebroplasty: how good is the evidence?. [Review] [67 refs]. *Current Rheumatology Reports* 2007; 9(1):57-65.

156. McKiernan, F., Jensen, R., Faciszewski, T. The dynamic mobility of vertebral compression fractures. *Journal of Bone & Mineral Research* 2003; 18(1):24-29.

157. Farcy, J.P., Weidenbaum, M., Glassman, S.D. Sagittal index in management of thoracolumbar burst fractures. *Spine (Phila Pa 1976)* 1990; 15(9):958-965.

158. Jiang, G., Eastell, R., Barrington, N.A., Ferrar, L. Comparison of methods for the visual identification of prevalent vertebral fracture in osteoporosis. *Osteoporosis International* 2004; 15:887-896.

159. Glassman, S.D., Bridwell, K., Dimar, J.R., Horton, W., Berven, S., Schwab, F. The Impact of Positive Sagittal Balance in Adult Spinal Deformity. *Spine* 2005; 30(18):2024-2029.

160. National Osteoporosis Foundation Working Group on Vertebral Fractures. Assessing vertebral fractures. *Journal of Bone & Mineral Research* 1995; 10:518-523.

161. Genant, H.K., Jergas, M. Assessment of prevalent and incident fractures in osteoporosis research. *Osteoporosis International* 2003; 14:S43-S55.

162. Grigoryan, M., Guermazi, A., Roemer, F.W., Delmas, P.D., Genant, H.K. Recognizing and reporting osteoporotic vertebral fractures. *European Spine Journal* 2003; 12:S104-S112.

163. Ferrar, L., Jiang, G., Adams, J., Eastell, R. Identification of vertebral fractures: an update. *Osteoporosis International* 2005; 16:717-728.

164. Fribourg, D., Tang, C., Sra, P., Delamarter, R., Bae, H. Incidence of subsequent vertebral fracture after kyphoplasty. *Spine* 2004; 29(20):2270-2276.

165. Klazen, C.A.H., Venmans, A., de Vries, J., van Rooij, W.J., Jansen, F.H., Blonk, M.C. et al. Percutaneous vertebroplasty is not a risk factor for new osteoporotic

compression fractures: results from VERTOS II. *American Journal of Neuroradiology* 2010; 31(8):1447-1450.

166. Mathis, J.M. Percutaneous vertebroplasty: complication avoidance and technique optimization. *American Journal of Neuroradiology* 2003; 24(8):1697-1706.

167. Venmans, A., Klazen, C.A., van Rooij, W.J., de Vries, J., Mali, W.P., Lohle, P.N. Postprocedural CT for perivertebral cement leakage in percutaneous vertebroplasty is not necessary - results from VERTOS II. *Neuroradiology* 2011; 53(1):19-22.

168. Jensen, M.E., Evans, A.J. Cardiovascular collapse and death during vertebroplasty - response. *Radiology* 2003; 228(3):902-903.

169. Jensen, M.E., Evans, A.J., Mathis, J.M., Kallmes, D.F., Cloft, H.J., Dion, J.E. Percutaneous polymethylmethacrylate vertebroplasty in the treatment of osteoporotic vertebral body compression fractures: technical aspects. *Ajnr: American Journal of Neuroradiology* 1997; 18(10):1897-1904.

170. Belinson, S.E. Percutaneous vertebroplasty or kyphoplasty for vertebral fractures caused by osteoporosis. *Tec Assessment Program* 2011; 25(9):1-56.

171. Martin, D.J., Rad, A.E., Kallmes, D.F. Prevalence of extravertebral cement leakage after vertebroplasty: procedural documentation versus CT detection. *Acta Radiol* 2012; 53(5):569-572.

172. Venmans, A., Klazen, C.A.H., Lohle, P.N.M., van Rooij, W.J., Verhaar, H.J.J., de Vries, J. et al. Percutaneous vertebroplasty and pulmonary cement embolism: results from VERTOS II. *American Journal of Neuroradiology* 2010; 31(8):1451-1453.

173. Klazen, C., Lohle, P., Jansen, F., Schoemaker, M., Elgersma, O., Van, E.K. et al. 1-year results of the VERTOS II trial: Vertebroplasty versus conservative therapy. *CardioVascular and Interventional Radiology* 2009; 32(Suppl 2):313.

174. Exponent. Mortality and complication risks for operated and non-operated vertebral fracture patients in the Medicare population. 2012.

175. Lange, A., Braun, S. Survival analysis using German claims data of patients with osteoporotic vertebral compression fractures (OVCF). Study Report. *Herescon Gmbh* 2012; January.

176. Loke, Y.K., Price, D., Herxheimer, A. Systematic reviews of adverse effects: framework for a structured approach. *BMC Med Res Methodol* 2007; 7:32.

177. Golder, S., Loke, Y., McIntosh, H.M. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. *BMC Med Res Methodol* 2006; 6:3.

178. Brewer, T., Colditz, G.A. Postmarketing surveillance and adverse drug reactions: Current perspectives and future needs. *JAMA* 1999; 281:824-829.

179. Cranney, A., Tugwell, P., Wells, G., Guyatt, G. Systematic reviews of randomized trials in osteoporosis: Introduction and methodology. *Endocrine Reviews* 2002; 23:497-507.

180. Mousavi, P., Roth, S., Finkelstein, J., Cheung, G., Whyne, C. Volumetric quantification of cement leakage following percutaneous vertebroplasty in metastatic and osteoporotic vertebrae. *Journal of Neurosurgery* 2003; 99(1:Suppl):Suppl-9.

181. Cotten, A., Dewatre, F., Cortet, B., Assaker, R., Leblond, D., Duquesnoy, B. et al. Percutaneous vertebroplasty for osteolytic metastases and myeloma: effects of the percentage of lesion filling and the leakage of methyl methacrylate at clinical follow-up. *Radiology* 1996; 200(2):525-530.

182. Review Manager (RevMan) Version 5.1. [Copenhagen: The Nordic Cochrane Centre; 2011.

183. Blasco, J., Garcia, A., Manzanera, L.S.R., MacHo, J.M., Peris, P., Jaume, P. et al. Randomized trial comparing vertebroplasty and conservative treatment analyzing pain relief and quality of life on the long term basis. *CardioVascular and Interventional Radiology* 2010; 33(Suppl 2):182-183.

184. Martinez-Ferrer, A., Blasco, J., Carrasco, J.L., Monegal, A., Pomes, J., Guaabens, N. et al. Effect of vertebroplasty on the quality of life of patients with pain related to osteoporotic vertebral fractures preliminary results of a randomized trial. *Bone* 2011; 48(Suppl 2):S161.

185. Rousing, R., Hansen, K.L., Andersen, M.O., Jespersen, S.M., Thomsen, K., Lauritsen, J.M. Twelve-months follow-up in forty-nine patients with acute/semiacute osteoporotic vertebral fractures treated conservatively or with percutaneous vertebroplasty. A clinical randomized study. *Spine* 2010; 35(5):478-482.

186. Muijs, S.P., Van Erkel, A.R., Dijkstra, P.D.S. Treatment of painful osteoporotic vertebral compression fractures: a brief review of the evidence for percutaneous vertebroplasty. *Journal of Bone & Joint Surgery - British Volume* 2011; 93-B(9):1149-1153.

187. Do, H.M., Marcellus, M.L., Weir, R.U., Marks, M.P. Percutaneous vertebroplasty versus medical therapy for treatment of acute vertebral body compression fractures: a prospective randomized study. *Internet* 2002; Available from http://members.asnr.org/abstracts/2002/02-O-1027-ASNR.pdf (accessed Feb. 2012).

188. Kallmes, D.E., Jensen, M.E., Marx, W.F., Sinha, R.S., Schweickert, P.A., Jarvik, J.G. A pilot study for a sham-controlled, randomized, prospective, crossover trial of percutaneous vertebroplasty. *Internet* 2002; Available from http://members.asnr.org/abstracts/2002/02-O-632-ASNR.pdf (accessed Feb. 2012).

189. ClinicalTrials.gov. KAVIAR study - Kyphoplasty And Vertebroplasty in the Augmentation and Restoration of vertebral body compression fractures. *Internet* 2012; Available from http://clinicaltrials.gov/ct2/show/NCT00323609?term=NCT00323609&rank=1 (accessed Feb. 2012).

190. Longo, U.G., Loppini, M., Denaro, L., Brandi, M.L., Maffulli, N., Denaro, V. The effectiveness and safety of vertebroplasty for osteoporotic vertebral compression fractures. A double blind, prospective, randomized, controlled study. *Clinical Cases in Mineral and Bone Metabolism* 2010; 7(2):109-113.

191. ClinicalTrials.gov. Percutaneous vertebroplasty versus conservative treatment of pain. *Internet* 2008; Available from http://clinicaltrials.gov/ct2/show/NCT00203554?term=00203554&rank=1 (accessed Sept. 2010).

192. ClinicalTrials.gov. Cost Effectiveness and Efficacy of kyphoplasty and vertebroplasty trial. *Internet* 2011; Available from http://clinicaltrials.gov/ct2/show/NCT00279877?term=NCT00279877&rank=1 (accessed Feb. 2012).

193. ClinicalTrials.gov. Comparison of balloon kyphoplasty, vertebroplasty and conservative management in acute osteoporotic vertebral fractures (OSTEO-6). *Internet* 2011; Available from http://clinicaltrials.gov/ct2/show/NCT00749060?term=NCT+00749060&rank=1 (accessed Sept. 2012).

194. ClinicalTrials.gov. Comparison of balloon kyphoplasty and vertebroplasty in subacute osteoporotic vertebral fractures (OSTEO+6). *Internet* 2011; Available from http://clinicaltrials.gov/ct2/show/NCT00749086?term=00749086&rank=1 (accessed Sept. 2011).

195. ClinicalTrials.gov. A trial of vertebroplasty for painful acute osteoporotic vertebral fractures (VERTOS IV). *Internet* 2011; Available from http://clinicaltrials.gov/ct2/show/NCT01200277?term=01200277&rank=1 (accessed Nov. 2011).

196. Boonen, S., Van, M.J., Bastian, L., Cummings, S.R., Ranstam, J., Tillman, J.B. et al. Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. *Journal of Bone & Mineral Research* 2011; 26(7):1627-1637.

197. Iranian Registry of Clinical Trials. The comparison of results of PV (percutaneous vertebroplasty) and conservative treatment in patients with osteoporotic VCF (vertebral compression fracture). *Internet* 2010; Available from http://www.irct.ir/searchresult.php?keyword=IRCT138804252193N1&id=2193&number=1&field=a&prt=1&total=1&m=1 (accessed Feb. 2012).

198. Higgins, J.P.T., Green, S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. *Internet* 2011; Available from http://www.cochrane-handbook.org/ (accessed Apr. 2012).

199. Kallmes, D.F., Comstock, B.A., Gray, L.A., Heagerty, P.J., Hollingworth, W., Turner, J.A. et al. Baseline pain and disability in the Investigational Vertebroplasty Efficacy and Safety Trial. *Ajnr: American Journal of Neuroradiology* 2009; 30(6):1203-1205.

200. Mathis, J.M., Ortiz, O., Zoarski, G.H. Vertebroplasty versus kyphoplasty: a comparison and contrast. *American Journal of Neuroradiology* 2004; 25(5):840-845.

201. Khosla, A., Turner, J.A., Jarvik, J.G., Gray, L.A., Kallmes, D.F. Impact of pain question modifiers on spine augmentation outcome. *Radiology* 2010; 257(2):477-482.

202. Orr, R.D. Vertebroplasty, cognitive dissonance, and evidence-based medicine: what do we do when the 'evidence' says we are wrong? *Cleveland Clinic Journal of Medicine* 2010; 77(1):8-11.

203. Van Meirhaeghe, J.K., Boonen, S., Bastian, L., Cummings, S., Ranstam, J., Tillman, J. et al. A randomized trial of balloon kyphoplasty and nonsurgical care for patients with acute vertebral compression fractures: Two year results. *Osteoporosis International* 2010; 21(Suppl 5):S667-S668.

204. ClinicalTrials.gov. FREE study - Fracture Reduction Evaluation. *Internet* 2010; Available from http://clinicaltrials.gov/ct2/show/NCT00211211?term=kyphoplasty&rank=17 (accessed Nov. 2011).

205. Blattert, T.R., Josten, C. Treatment of osteoporotic vertebral body fractures by means of percutaneous Balloon kyphoplasty: Long term results of a prospective, clinical trial. *Osteoporosis International* 2010; 21(Suppl 1):S384.

206. Diel, P., Reuss, W., Aghayev, E., Moulin, P., Roder, C., SWISSspine Registry Group. SWISSspine-a nationwide health technology assessment registry for balloon kyphoplasty: methodology and first results. *Spine Journal: Official Journal of the North American Spine Society* 2010; 10(11):961-971.

207. Majd, M.E., Farley, S., Holt, R.T. Preliminary outcomes and efficacy of the first 360 consecutive kyphoplasties for the treatment of painful osteoporotic vertebral compression fractures. *Spine Journal: Official Journal of the North American Spine Society* 2005; 5(3):244-255.

208. Alvarez, L., Perez-Higueras, A., Granizo, J.J., de, M., I, Quinones, D., Rossi, R.E. Predictors of outcomes of percutaneous vertebroplasty for osteoporotic vertebral fractures. *Spine* 2005; 30(1):87-92.

209. Diel, P., Merky, D., Roder, C., Popp, A., Perler, M., Heini, P.F. Safety and efficacy of vertebroplasty: Early results of a prospective one-year case series of osteoporosis patients in an academic high-volume center. *Indian Journal of Orthopaedics* 2009; 43(3):228-233.

210. Evans, A.J., Jensen, M.E., Kip, K.E., DeNardo, A.J., Lawler, G.J., Negin, G.A. et al. Vertebral compression fractures: pain reduction and improvement in functional mobility after percutaneous polymethylmethacrylate vertebroplasty retrospective report of 245 cases. *Radiology* 2003; 226(2):366-372.

211. Lee, S.T., Chen, J.F. Closed reduction vertebroplasty for the treatment of osteoporotic vertebral compression fractures. Technical note. *Journal of Neurosurgery* 2004; 100(4:Suppl Spine):Suppl-6.

212. Masala, S., Mammucari, M., Angelopoulos, G., Fiori, R., Massari, F., Faria, S. et al. Percutaneous vertebroplasty in the management of vertebral osteoporotic fractures. Short-term, mid-term and long-term follow-up of 285 patients. *Skeletal Radiology* 2009; 38(9):863-869.

213. Mpotsaris, A., Abdolvahabi, R., Hoffleith, B., Nickel, J., Harati, A., Loehr, C. et al. Percutaneous vertebroplasty in vertebral compression fractures of benign or malignant origin: a prospective study of 1188 patients with follow-up of 12 months. *Deutsches Arzteblatt International* 2011; 108(19):331-338.

214. Ryu, K.S., Park, C.K. The prognostic factors influencing on the therapeutic effect of percutaneous vertebroplasty in treating osteoporotic vertebral compression fractures. *Journal of Korean Neurosurgical Society* 2009; 45(1):16-23.

215. Edidin, A.A., Ong, K.L., Lau, E., Kurtz, S.M. Mortality risk for operated and nonoperated vertebral fracture patients in the medicare population. *Journal of Bone & Mineral Research* 2011; 26(7):1617-1626.

216. Lange, A., Braun, S. Survival analysis using German claims data of patients with Osteoporotic Vertebral Compression Fractures (OVCF). Propensity Score Matching Results (BKP vs. VP). *Herescon Gmbh* 2012; February.

217. Padovani, B., Kasriel, O., Brunner, P., Peretti-Viton, P. Pulmonary embolism caused by acrylic cement: a rare complication of percutaneous vertebroplasty. *Ajnr: American Journal of Neuroradiology* 1999; 20(3):375-377.

218. Masala, S., Mastrangeli, R., Petrella, M.C., Massari, F., Ursone, A., Simonetti, G. Percutaneous vertebroplasty in 1,253 levels: results and long-term effectiveness in a single centre. *European Radiology* 2009; 19(1):165-171.

219. Saliou, G., Rutgers, D.R., Kocheida, E.M., Langman, G., Meurin, A., Deramond, H. et al. Balloon-related complications and technical failures in kyphoplasty for vertebral fractures. *Ajnr: American Journal of Neuroradiology* 2010; 31(1):175-179.

220. Harrop, J.S., Prpa, B., Reinhardt, M.K., Lieberman, I. Primary and secondary osteoporosis' incidence of subsequent vertebral compression fractures after kyphoplasty. *Spine* 2004; 29(19):2120-2125.

221. Kulcsar, Z., Marosfoi, M., Berentei, Z., Szikora, I. Frequency and risk factors of adjacent fractures after vertebroplasty. *Neuroradiology* 2009; 51(Suppl 1):S93.

222. Tseng, Y.Y., Yang, T.C., Tu, P.H., Lo, Y.L., Yang, S.T. Repeated and multiple new vertebral compression fractures after percutaneous transpedicular vertebroplasty. *Spine* 2009; 34(18):1917-1922.

223. Uppin, A.A., Hirsch, J.A., Centenera, L.V., Pfiefer, B.A., Pazianos, A.G., Choi, I.S. Occurrence of new vertebral body fracture after percutaneous vertebroplasty in patients with osteoporosis. *Radiology* 2003; 226(1):119-124.

224. Donovan, M.A., Khandji, A.G., Siris, E. Multiple adjacent vertebral fractures after kyphoplasty in a patient with steroid-induced osteoporosis. *Journal of Bone & Mineral Research* 2004; 19(5):712-713.

225. Han, I.H., Chin, D.K., Kuh, S.U., Kim, K.S., Jin, B.H., Yoon, Y.S. et al. Magnetic resonance imaging findings of subsequent fractures after vertebroplasty. *Neurosurgery* 744; 64(4):740-744.

226. Hierholzer, J., Fuchs, H., Westphalen, K., Baumann, C., Slotosch, C., Schulz, R. Incidence of symptomatic vertebral fractures in patients after percutaneous vertebroplasty. *Cardiovascular & Interventional Radiology* 2008; 31(6):1178-1183.

227. Kim, H.S., Jeon, K.H., Choi, W.J., Kim, K.T., Ju, C.I., Kim, S.W. et al. Spinal instability predicting score (SIPS) for following fractures after vertebroplasty in patients with osteoporotic vertebral compression fractures. *Spine Journal* 2010; 10(9 Suppl):106S.

228. Lee, W.S., Sung, K.H., Jeong, H.T., Sung, Y.S., Hyun, Y.I., Choi, J.Y. et al. Risk factors of developing new symptomatic vertebral compression fractures after percutaneous vertebroplasty in osteoporotic patients. *European Spine Journal* 2006; 15(12):1777-1783.

229. Lo, Y.P., Chen, W.J., Chen, L.H., Lai, P.L. New vertebral fracture after vertebroplasty. *Journal of Trauma-Injury Infection & Critical Care* 2008; 65(6):1439-1445.

230. Nene, A. Vertebroplasty-happily ever after? *Osteoporosis International* 2010; 21(Suppl 5):S756.

231. Nunley, P.D., Jawahar, A. A Comparison of Clinical Outcomes and Adjacent Level Fractures in Patients Receiving Vertebroplasty for Osteoporotic Compression Fractures Using Cortoss or Poly Methyl Methacrylate: Prospective, Randomized Trial. *Neurosurgery* 2009; 65(2):409.

232. Syed, M.I., Patel, N.A., Jan, S., Harron, M.S., Morar, K., Shaikh, A. New symptomatic vertebral compression fractures within a year following vertebroplasty in osteoporotic women. *Ajnr: American Journal of Neuroradiology* 2005; 26(6):1601-1604.

233. Tatsumi, R.L., Ching, A.C., Byrd, G.D., Hiratzka, J.R., Threlkeld, J.E., Hart, R.A. Predictors and prevalence of patients undergoing additional kyphoplasty procedures after an initial kyphoplasty procedure. *Spine Journal: Official Journal of the North American Spine Society* 2010; 10(11):979-986.

234. Trout, A.T., Kallmes, D.F., Kaufmann, T.J. New fractures after vertebroplasty: adjacent fractures occur significantly sooner. *Ajnr: American Journal of Neuroradiology* 2006; 27(1):217-223.

235. Trout, A.T., Kallmes, D.F., Lane, J.I., Layton, K.F., Marx, W.F. Subsequent vertebral fractures after vertebroplasty: association with intraosseous clefts. *Ajnr: American Journal of Neuroradiology* 2006; 27(7):1586-1591.

236. Trout, A.T., Kallmes, D.F. Does vertebroplasty cause incident vertebral fractures? A review of available data.. *Ajnr: American Journal of Neuroradiology* 2006; 27(7):1397-1403.

237. Mudano, A.S., Bian, J., Cope, J.U., Curtis, J.R., Gross, T.P., Allison, J.J. et al. Vertebroplasty and kyphoplasty are associated with an increased risk of secondary vertebral compression fractures: a population-based cohort study. *Osteoporosis International* 2009; 20(5):819-826.

238. Lin, C.C., Yen, P.S., Wen, S.H. Fluid sign in the treated bodies after percutaneous vertebroplasty. *Neuroradiology* 2008; 50(11):955-961.

239. Baroud, G., Bohner, M. Biomechanical impact of vertebroplasty. Postoperative biomechanics of vertebroplasty. *Joint, Bone, Spine: Revue Du Rhumatisme* 2006; 73(2):144-150.

240. Berlemann, U., Ferguson, S.J., Nolte, L.P., Heini, P.F. Adjacent vertebral failure after vertebroplasty. A biomechanical investigation. *Journal of Bone & Joint Surgery - British Volume* 2002; 84(5):748-752.

241. Baroud, G., Nemes, J., Heini, P., Steffen, T. Load shift of the intervertebral disc after a vertebroplasty: a finite-element study. *European Spine Journal* 2003; 12(4):421-426.

242. Baroud, G., Heini, P., Nemes, J., Bohner, M., Ferguson, S., Steffen, T. Biomechanical explanation of adjacent fractures following vertebroplasty. *Radiology* 2003; 229(2):606-607.

243. Farooq, N., Park, J.C., Pollintine, P., Annesley-Williams, D.J., Dolan, P. Can vertebroplasty restore normal load-bearing to fractured vertebrae? *Spine* 2005; 30(15):1723-1730.

244. Yang, S.C., Chen, W.J., Yu, S.W., Tu, Y.K., Kao, Y.H., Chung, K.C. Revision strategies for complications and failure of vertebroplasties. *European Spine Journal* 2008; 17(7):982-988.

245. Abdul-Jalil, Y., Bartels, J., Alberti, O., Becker, R. Delayed presentation of pulmonary polymethylmethacrylate emboli after percutaneous vertebroplasty. *Spine* 2007; 32(20):E589-E593.

246. Agko, M., Nazzal, M., Jamil, T., Castillo-Sang, M., Clark, P., Kasper, G. Prevention of cardiopulmonary embolization of polymethylmethacrylate cement fragment after kyphoplasty with insertion of inferior vena cava filter. *Journal of Vascular Surgery* 2010; 51(1):210-213.

247. Baumann, A., Tauss, J., Baumann, G., Tomka, M., Hessinger, M., Tiesenhausen, K. Cement embolization into the vena cava and pulmonal arteries after vertebroplasty: interdisciplinary management. *European Journal of Vascular & Endovascular Surgery* 2006; 31(5):558-561.

248. Bernhard, J., Heini, P.F., Villiger, P.M. Asymptomatic diffuse pulmonary embolism caused by acrylic cement: an unusual complication of percutaneous vertebroplasty. *Annals of the Rheumatic Diseases* 2003; 62(1):85-86.

249. Biega, T.J., Lettieri, C.J., Levy, L.M., Venbrux, A.C. Linear pulmonary opacities in an asymptomatic patient. *Respiration* 2006; 73(5):705-707.

250. Bonardel, G., Pouit, B., Gontier, E., Dutertre, G., Mantzarides, M., Goasguen, O. et al. Pulmonary cement embolism after percutaneous vertebroplasty: a rare and nonthrombotic cause of pulmonary embolism. *Clinical Nuclear Medicine* 2007; 32(8):603-606.

251. Cadeddu, C., Nocco, S., Secci, E., Deidda, M., Pirisi, R., Mercuro, G. Echocardiographic accidental finding of asymptomatic cardiac and pulmonary embolism caused by cement leakage after percutaneous vertebroplasty. *European Journal of Echocardiography* 2009; 10(4):590-592.

252. Caynak, B., Onan, B., Sagbas, E., Duran, C., Akpinar, B. Cardiac tamponade and pulmonary embolism as a complication of percutaneous vertebroplasty.. *Annals of Thoracic Surgery* 2009; 87(1):299-301.

253. Chen, H.L., Wong, C.S., Ho, S.T., Chang, F.L., Hsu, C.H., Wu, C.T. A lethal pulmonary embolism during percutaneous vertebroplasty. *Anesthesia & Analgesia* 2002; 95(4):1060-1062.

254. Dastidar, J. History of PE/DVT + vertebroplasty + dyspnea = ? *Journal of Hospital Medicine* 2011; Conference(var.pagings):S170.

255. Finch, L., Cheng, S.G., Steinberg, K.P., Stern, E.J. Polymethylmethacrylate pulmonary emboli. *Clinical Pulmonary Medicine* 2002; 9(2):133-134.

256. Francois, K., Taeymans, Y., Poffyn, B., Van, N.G. Successful management of a large pulmonary cement embolus after percutaneous vertebroplasty: a case report. *Spine* 2003; 28(20):E424-E425.

257. Freitag, M., Gottschalk, A., Schuster, M., Wenk, W., Wiesner, L., Standl, T.G. Pulmonary embolism caused by polymethylmethacrylate during percutaneous vertebroplasty in orthopaedic surgery. *Acta Anaesthesiologica Scandinavica* 2006; 50(2):248-251.

258. Grahe, J.S., Casey, L., White, G. Pulmonary cement embolism following percutaneous vertebroplasty. *Chest* 2004; 126(4):956S.

259. Harris, B., Briggs, G., Dennis, C. Cement pulmonary embolism as a consequence of vertebroplasty. *Internal Medicine Journal* 2007; 37(3):196-197.

260. Jang, J.S., Lee, S.H., Jung, S.K. Pulmonary embolism of polymethylmethacrylate after percutaneous vertebroplasty: a report of three cases. *Spine* 2002; 27(19):E416-E418.

261. Kim, Y.J., Lee, J.W., Park, K.W., Yeom, J.S., Jeong, H.S., Park, J.M. et al. Pulmonary cement embolism after percutaneous vertebroplasty in osteoporotic vertebral compression fractures: incidence, characteristics, and risk factors. *Radiology* 2009; 251(1):250-259.

262. Kovalenko, B., Rao, Q., Hughes, T. Pulmonary embolism caused by acrylic cement from prior vertebroplasty. *CardioVascular and Interventional Radiology* 2009; 32(Suppl 2):334.

263. Sang, E.L., Chang, S.A., Kim, M.S., Kim, S.Y., Han, J.K., Jang, H.J. et al. Acrylic cement foreign body and thrombus in right atrium causing pulmonary embolism after percutaneous vertebroplasty. *Korean Circulation Journal* 2006; 36(10):713-715.

264. Leroux, G., Costedoat-Chalumeau, N., Chiras, J., de Gennes, C., Piette, J.C. A vertebroplasty with dyspnea. *Revue De Medecine Interne* 2007; 28(7):492-494.

265. Liliang, P.C., Lu, K., Liang, C.L., Tsai, Y.D., Hsieh, C.H., Chen, H.J. Dyspnoea and chest pain associated with pulmonary polymethylmethacrylate embolism after percutaneous vertebroplasty. *Injury* 2007; 38(2):245-248.

266. Lim, K.J., Yoon, S.Z., Jeon, Y.S., Bahk, J.H., Kim, C.S., Lee, J.H. et al. An intraatrial thrombus and pulmonary thromboembolism as a late complication of percutaneous vertebroplasty. *Anesthesia & Analgesia* 2007; 104(4):924-926.

267. Lim, S.H., Kim, H., Kim, H.K., Baek, M.J. Multiple cardiac perforations and pulmonary embolism caused by cement leakage after percutaneous vertebroplasty. *European Journal of Cardio-Thoracic Surgery* 2008; 33(3):510-512.

268. MacTaggart, J.N., Pipinos, I.I., Johanning, J.M., Lynch, T.G. Acrylic cement pulmonary embolus masquerading as an embolized central venous catheter fragment.. *Journal of Vascular Surgery* 2006; 43(1):180-183.

269. Moll, S., Kuzma, C. Images in vascular medicine: cement pulmonary embolism. *Vascular Medicine* 2010; 15(4):339-340.

270. Monticelli, F., Meyer, H.J., Tutsch-Bauer, E. Fatal pulmonary cement embolism following percutaneous vertebroplasty (PVP). *Forensic Science International* 2005; 149(1):35-38.

271. Moon, S., Lee, S., Kong, G., Kim, J., Lee, E. Large pulmonary embolus after percutaneous vertebroplasty in osteoporotic compression fracture - A case report. *Bone* 2009; 44(Suppl 1):S86.

272. Muller, M., Biedermann, M., Strecker, W. A complication during kyphoplasty. Cement penetration through the azygos vein into the superior vena cava. *Orthopade* 2006; 35(11):1183-1186.

273. Neuwirth, J., Weber, J.C., Kohler, B. Pulmonary cement embolism after vertebroplasty in multiple osteoporotic vertebral fractures. *Deutsche Medizinische Wochenschrift* 2006; 131:2275-2276.

274. Perrin, C., Jullien, V., Padovani, B., Blaive, B. Percutaneous vertebroplasty complicated by pulmonary embolus of acrylic cement. *Revue Des Maladies Respiratoires* 1999; 16(2):215-217.

275. Pleser, M., Roth, R., Worsdorfer, O., Manke, C. Pulmonary embolism caused by PMMA in percutaneous vertebroplasty. Case report and review of the literature. *Unfallchirurg* 2004; 107(9):807-811.

276. Pott, L., Wippermann, B., Hussein, S., Gunther, T., Brusch, U., Fremerey, R. PMMA pulmonary embolism and post interventional associated fractures after percutaneous vertebroplasty. *Orthopade* 2005; 34(7):698-702.

277. Quesada, N., Mutlu, G.M. Images in cardiovascular medicine. Pulmonary embolization of acrylic cement during vertebroplasty. *Circulation* 2006; 113(8):e295-e296.

278. Radcliff, K.E., Reitman, C.A., Delasotta, L.A., Hong, J., DiIorio, T., Zaslavsky, J. et al. Pulmonary cement embolization after kyphoplasty: a case report and review of the literature. *Spine Journal: Official Journal of the North American Spine Society* 2010; 10(10):e1-e5.

279. Righini, M., Sekoranja, L., Le, G.G., Favre, I., Bounameaux, H., Janssens, J.P. Pulmonary cement embolism after vertebroplasty. *Thrombosis & Haemostasis* 2006; 95(2):388-389.

280. Schneider, L., Plit, M. Pulmonary embolization of acrylic cement during percutaneous vertebroplasty. *Internal Medicine Journal* 2007; 37(6):423-4U8.

281. Schoenes, B., Bremerich, D.H., Risteski, P.S., Thalhammer, A., Meininger, D. Cardiac perforation after vertebroplasty. *Anaesthesist* 2008; 57(2):147-150.

282. Scroop, R., Eskridge, J., Britz, G.W. Paradoxical cerebral arterial embolization of cement during intraoperative vertebroplasty: case report.[Retraction in Eskridge J, Britz GW. AJNR Am J Neuroradiol. 2004 Mar;25(3):B1; PMID: 15043050]. *Ajnr: American Journal of Neuroradiology* 2002; 23(5):868-870.

283. Seo, J.S., Kim, Y.J., Choi, B.W., Kim, T.H., Choe, K.O. MDCT of pulmonary embolism after percutaneous vertebroplasty. *AJR American Journal of Roentgenology* 2005; 184(4):1364-1365.

284. Shalshin, A., Brar, N., Chawla, S., Islam, T. Pulmonary cement embolism: Rare complication of Kyphoplasty. *Chest* 2004; 138(4):13A.

285. Son, K.H., Chung, J.H., Sun, K., Son, H.S. Cardiac perforation and tricuspid regurgitation as a complication of percutaneous vertebroplasty. *European Journal of Cardio-Thoracic Surgery* 2008; 33(3):508-509.

286. Stricker, K., Orler, R., Yen, K., Takala, J., Luginbuhl, M. Severe hypercapnia due to pulmonary embolism of polymethylmethacrylate during vertebroplasty. *Anesthesia & Analgesia* 1000; 98(4):1184-1186.

287. Torres Machi, M.L., Suarez, R.V., Medina, R.C., Gil, B.F., Ojeda, B.N., Rodriguez-Perez, A. Pulmonary embolism caused by cement following vertebroplasty. *Revista Espanola Anestesiologica Reanim* 2003; 50:489-491.

288. Tozzi, P., Abdelmoumene, Y., Corno, A.F., Gersbach, P.A., Hoogewoud, H.M., von Segesser, L.K. Management of pulmonary embolism during acrylic vertebroplasty. *Annals of Thoracic Surgery* 2002; 74(5):1706-1708.

289. Yoo, K.Y., Jeong, S.W., Yoon, W., Lee, J. Acute respiratory distress syndrome associated with pulmonary cement embolism following percutaneous vertebroplasty with polymethylmethacrylate. *Spine* 2004; 29(14):E294-E297.

290. Zaccheo, M.V., Rowane, J.E., Costello, E.M. Acute respiratory failure associated with polymethyl methacrylate pulmonary emboli after percutaneous vertebroplasty. *American Journal of Emergency Medicine* 2008; 26(5):636-637.

291. Gaye, M., Fuentes, S., Pech-Gourg, G., Benhima, Y., Dufour, H. [Spondylitis following vertebroplasty. Case report and review of the literature] [French]. *Neurochirurgie* 2008; 54(4):551-555.

292. Ivo, R., Sobottke, R., Seifert, H., Ortmann, M., Eysel, P. Tuberculous spondylitis and paravertebral abscess formation after kyphoplasty: a case report. *Spine* 2010; 35(12):E559-E563.

293. Lee, C.B., Kim, H.S., Kim, Y.J. Pyogenic spondylitis after vertebroplasty - a report of two cases -. *Asian Spine Journal* 2007; 1(2):106-109.

294. Mummaneni, P.V., Walker, D.H., Mizuno, J., Rodts, G.E. Infected vertebroplasty requiring 360 degrees spinal reconstruction: long-term follow-up review. Report of two cases. *Journal of Neurosurgery Spine* 2006; 5(1):86-89.

295. Schmid, K.E., Boszczyk, B.M., Bierschneider, M., Zarfl, A., Robert, B., Jaksche, H. Spondylitis following vertebroplasty: a case report. *European Spine Journal* 2005; 14(9):895-899.

296. Schofer, M.D., Lakemeier, S., Peterlein, C.D., Heyse, T.J., Quante, M. Primary pyogenic spondylitis following kyphoplasty: a case report. *Journal of Medical Case Reports [Electronic Resource]* 2011; 5:101.

297. Shin, J.H., Ha, K.Y., Kim, K.W., Lee, J.S., Joo, M.W. Surgical treatment for delayed pyogenic spondylitis after percutaneous vertebroplasty and kyphoplasty. Report of 4 cases.. *Journal of Neurosurgery Spine* 2008; 9(3):265-272.

298. Syed, M.I., Avutu, B., Shaikh, A., Sparks, H., Mohammed, M.I., Morar, K. Vertebral osteomyelitis following vertebroplasty: is acne a potential contraindication and are prophylactic antibiotics mandatory prior to vertebroplasty? *Pain Physician* 2009; 12(4):E285-E290.

299. Vats, H.S., McKiernan, F.E. Infected vertebroplasty: case report and review of literature.. *Spine* 2006; 31(22):E859-E862.

300. Walker, D.H., Mummaneni, P., Rodts, G.E., Jr. Infected vertebroplasty. Report of two cases and review of the literature. *Neurosurgical Focus* 2004; 17(6):E6.

301. Wendling, D., Runge, M., Toussirot, E., Bertolini, E., Prati, C. Vertebral osteitis adjacent to kyphoplasty. *Joint, Bone, Spine: Revue Du Rhumatisme* 2010; 77(1):67-69.

302. Yu, S.W., Chen, W.J., Lin, W.C., Chen, Y.J., Tu, Y.K. Serious pyogenic spondylitis following vertebroplasty--a case report. *Spine* 2004; 29(10):E209-E211.

303. Farahvar, A., Dubensky, D., Bakos, R. Perforation of the right cardiac ventricular wall by polymethylmethacrylate after lumbar kyphoplasty. *Journal of Neurosurgery Spine* 2009; 11(4):487-491.

304. Kim, S.Y., Seo, J.B., Do, K.H., Lee, J.S., Song, K.S., Lim, T.H. Cardiac perforation caused by acrylic cement: a rare complication of percutaneous vertebroplasty. *AJR American Journal of Roentgenology* 2005; 185(5):1245-1247.

305. Kao, F.C., Tu, Y.K., Lai, P.L., Yu, S.W., Yen, C.Y., Chou, M.C. Inferior vena cava syndrome following percutaneous vertebroplasty with polymethylmethacrylate. *Spine* 2008; 33(10):E329-E333.

306. White, J.B., Thielen, K.R., Kallmes, D.F. Putative risk of substantial venous air embolism during vertebroplasty: a technical observation. *Spine* 2009; 34(14):1526-1528.

307. Dash, A., Brinster, D.R. Open heart surgery for removal of polymethylmethacrylate after percutaneous vertebroplasty. *Annals of Thoracic Surgery* 2011; 91(1):276-278.

308. Park, J.H., Choo, S.J., Park, S.W. Images in cardiovascular medicine. Acute pericarditis caused by acrylic bone cement after percutaneous vertebroplasty. *Circulation* 2005; 111(6):e98.

309. Puri, A.S., Colen, R.R., Reddy, A.S., Groff, M.W., DiNobile, D., Killoran, T. et al. Lumbar artery pseudoaneurysm after percutaneous vertebroplasty: a unique vascular complication. *Journal of Neurosurgery Spine* 2011; 14(2):296-299.

310. Marden, F.A., Putman, C.M. Cement-embolic stroke associated with vertebroplasty. *Ajnr: American Journal of Neuroradiology* 2008; 29(10):1986-1988.

311. Biafora, S.J., Mardjetko, S.M., Butler, J.P., McCarthy, P.L., Gleason, T.F. Arterial injury following percutaneous vertebral augmentation: a case report. *Spine* 2006; 31(3):E84-E87.

312. Heo, D.H., Cho, Y.J. Segmental artery injury following percutaneous vertebroplasty using extrapedicular approach. *Journal of Korean Neurosurgical Society* 2011; 49(2):131-133.

313. Hard, J.M., Gonda, R.L., Kadakia, S.R. A novel approach to treatment of unexpected vertebroplasty complication. *Cardiovascular & Interventional Radiology* 2008; 31(6):1249-1251.

314. Ozturk, C., Tas, I., Hepguler, S., Dusunceli, Y. Paraplegia as a complication of vertebroplasty: a case report. *Osteoporosis International* 2007; 18(Suppl 1):S40.

315. Lee, B.J., Lee, S.R., Yoo, T.Y. Paraplegia as a complication of percutaneous vertebroplasty with polymethylmethacrylate: a case report.. *Spine* 2002; 27(19):E419-E422.

316. Birkenmaier, C., Seitz, S., Wegener, B., Glaser, C., Ruge, M.I., von, L.A. et al. Acute paraplegia after vertebroplasty caused by epidural hemorrhage. A case report. *Journal of Bone & Joint Surgery - American Volume* 2007; 89(8):1827-1831.

317. Lopes, N.M., Lopes, V.K. Paraplegia complicating percutaneous vertebroplasty for osteoporotic vertebral fracture: case report. *Arquivos De Neuro-Psiquiatria* 2004; 62(3B):879-881.

318. Lim, J.B., Park, J.S., Kim, E. Nonaneurysmal subarachnoid hemorrhage : rare complication of vertebroplasty. *Journal of Korean Neurosurgical Society* 2009; 45(6):386-389.

319. Heo, D.H., Cho, S.M., Cho, Y.J., Cho, J.H., Sheen, S.H. Heterotopic ossifications after vertebroplasty using calcium phosphate in osteoporotic vertebral compression fractures: Report of 2 cases. *World Neurosurgery* 1000; 73(3):207-209.

320. Nemeth, A.J., Lie-Nemeth, T.J., Marota, J.J., Pryor, J.C., Rabinov, J.D., Hirsch, J.A. Vertebral augmentation complicated by perioperative addisonian crisis. *Pain Physician* 2006; 9(3):257-260.

321. Sonmez, E., Yilmaz, C., Caner, H. Development of lumbar disc herniation following percutaneous vertebroplasty.. *Spine* 2010; 35(3):E93-E95.

322. Soyuncu, Y., Ozdemir, H., Soyuncu, S., Bigat, Z., Gur, S. Posterior spinal epidural abscess: an unusual complication of vertebroplasty. *Joint, Bone, Spine: Revue Du Rhumatisme* 2006; 73(6):753-755.

323. Syed, M.I., Jan, S., Patel, N.A., Shaikh, A., Marsh, R.A., Stewart, R.V. Fatal fat embolism after vertebroplasty: identification of the high-risk patient. *Ajnr: American Journal of Neuroradiology* 2006; 27(2):343-345.

324. Yang, C.T., Hou, S.M., Hou, C.H., Lin, F.L., Lin, C.C., Yang, R.S. Balloon kyphoplasty and vertebroplasty in the management of vertebral compression fracture: Does complication rate differ in countries or specialties of operators? *Osteoporosis International* 2010; 21(Suppl 5):S767-S768.

325. Perisinakis, K., Damilakis, J., Theocharopoulos, N., Papadokostakis, G., Hadjipavlou, A., Gourtsoyiannis, N. Patient exposure and associated radiation risks from fluoroscopically guided vertebroplasty or kyphoplasty. *Radiology* 2004; 232(3):701-707.

326. Fitousi, N.T., Efstathopoulos, E.P., Delis, H.B., Kottou, S., Kelekis, A.D., Panayiotakis, G.S. Patient and staff dosimetry in vertebroplasty. *Spine* 2006; 31(23):E884-E889.

327. Slipman, C.W., Lipetz, J.S., Jackson, H.B., Vresilovic, E.J. Deep venous thrombosis and pulmonary embolism as a complication of bed rest for low back pain. *Archives of Physical Medicine and Rehabilitation* 2000; 81(1):127-129.

328. Convertino, V.A., Bloomfield, S.A., Greenleaf, J.E. An overview of the issues: physiological effects of bed rest and restricted physical activity. *Medicine & Science in Sports & Exercise* 1997; 29(2).

329. Munk, P.L., Liu, D.M., Murphy, K.P., Baerlocher, M.O. Effectiveness of Vertebroplasty: A Recent Controversy. *Canadian Association of Radiologists*

Journal-Journal De l Association Canadienne Des Radiologistes 2009; 60(4):170-171.

330. Kaufmann, T.J., Trout, A.T., Kallmes, D.F. The effects of cement volume on clinical outcomes of percutaneous vertebroplasty. *Ajnr: American Journal of Neuroradiology* 2006; 27(9):1933-1937.

331. Al-Ali, F., Barrow, T., Luke, K. Vertebroplasty: what is important and what is not. *American Journal of Neuroradiology* 2009; 30(10):1835-1839.

332. Brinjikji, W., Comstock, B.A., Gray, L., Kallmes, D.F. Local Anesthesia with Bupivacaine and Lidocaine for vertebral fracture trial (LABEL): a report of outcomes and comparison with the Investigational Vertebroplasty Efficacy and Safety Trial (INVEST). *American Journal of Neuroradiology* 2010; 31(9):1631-1634.

333. Miller, F.G., Kallmes, D.F., Buchbinder, R. Vertebroplasty and the placebo response. *Radiology* 2011; 259:621-625.

334. Lotz, J.C. Trials of veretebroplasty for vertebral fractures. *New England Journal of Medicine* 2009; 361(21):2098.

335. Brinjikji, W., Comstock, B.A., Heagerty, P.J., Jarvik, J.G., Kallmes, D.F. Investigational Vertebroplasty Efficacy and Safety Trial: detailed analysis of blinding efficacy. *Radiology* 2010; 257(1):219-225.

336. Kallmes, D.F., Heagerty, P.J., Jarvik, J.G. Trials of vertebroplasty for vertebral fractures. *New England Journal of Medicine* 2009; 361(21):2099-2100.

337. Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Jüni, P., Altman, D.G. et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008; 336:601-605.

338. Psaty, B.M., Prentice, R.L. Minimizing bias in randomized trials: the importance of blinding. *JAMA* 2010; 304(7):793-794.

339. Brody, H.B., Brody, D. Placebo and health--II. Three perspectives on the placebo response: expectancy, conditioning, and meaning. *Adv Mind Body Med* 2000; 16(3):216-232.

340. Scott, D.J., Stohler, C.S., Egnatuk, C.M., Wang, H., Koeppe, R.A., Zubieta, J.K. Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron* 2007; 55(2):325-336.

341. McDonald, R.J., Gray, L.A., Cloft, H.J., Thielen, K.R., Kallmes, D.F. The effect of operator variability and experience in vertebroplasty outcomes. *Radiology* 2009; 253(2):478-485.

342. Syed, M., Shaikh, A., Akhter, T., Morar, K., Cortoss, S. Does age of fracture affect the outcome of vertebroplasty? Results from a prospective multicenter FDA-IDE study. *Journal of Vascular and Interventional Radiology* 2011; 22(3 Suppl):S86.

343. Blattert, T.R., Jestaedt, L., Weckbach, A. Suitability of a calcium phosphate cement in osteoporotic vertebral body fracture augmentation: a controlled, randomized, clinical trial of balloon kyphoplasty comparing calcium phosphate versus polymethylmethacrylate. *Spine* 2009; 34(2):108-114.

344. Anselmetti, G.C., Zoarski, G., Manca, A., Masala, S., Eminefendic, H., Russo, F. et al. Percutaneous vertebroplasty and bone cement leakage: clinical experience with a new high-viscosity bone cement and delivery system for vertebral augmentation in benign and malignant compression fractures. *Cardiovascular & Interventional Radiology* 2008; 31(5):937-947.

345. Yang, H.L., Zhao, L., Liu, J., Sanford, C.G., Jr., Chen, L., Tang, T. et al. Changes of pulmonary function for patients with osteoporotic vertebral compression fractures after kyphoplasty. *Journal of Spinal Disorders & Techniques* 2007; 20(3):221-225.

346. Alvarez, L., Alcaraz, M., Perez-Higueras, A., Granizo, J.J., de, M., I, Rossi, R.E. et al. Percutaneous vertebroplasty: functional improvement in patients with osteoporotic compression fractures. *Spine* 2006; 31(10):1113-1118.

347. Clark, W., Lyon, S., Burnes, J. Trials of vertebroplasty for vertebral fractures. *New England Journal of Medicine* 2009; 361(21):2097-2098.

348. Buchbinder, R., Osborne, R., Staples, M. Trials of vertebroplasty for vertical fractures. *New England Journal of Medicine* 2009; 361(21):2099.

349. Strom, O., Leonard, C., Marsh, D., Cooper, C. Cost-effectiveness of balloon kyphoplasty in patients with symptomatic vertebral compression fractures in a UK setting. *Osteoporosis International* 2010; 21(9):1599-1608.

350. Chen, C., Chen, L., Gu, Y., Xu, Y., Liu, Y., Bai, X. et al. Kyphoplasty for chronic painful osteoporotic vertebral compression fractures via unipedicular versus bipedicular approachment: a comparative study in early stage. *Injury* 2010; 41(4):356-359.

351. Chen, L., Yang, H., Tang, T. Unilateral versus bilateral balloon kyphoplasty for multilevel osteoporotic vertebral compression fractures: a prospective study. *Spine* 2011; 36(7):534-540.

352. Liu, J.T., Liao, W.J., Tan, W.C., Lee, J.K., Liu, C.H. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. *Osteoporos Int* 2009; 21(2):359-364.

353. Department of Health. Confirmation of Payment by Results (PbR) arrangements for 2011-12. 2012; available from http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_124356

354. Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). 2012; available from www.mortality.org

355. Ara, R., Brazier, J.E. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2010; 13[5], 509-518.

356. National Institute for Health and Clinical Excellence. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fracture. Final scope. 2011.

357. Barendregt, J.J. The Half-Cycle Correction: Banish Rather Than Explain It. *Medical Decision Making* 2009; 29(4):500-502.

358. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008; available from http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

359. Stevenson, M., Lloyd-Jones, M., Papaioannou, D., Stevenson, M., Lloyd-Jones, M., Papaioannou, D. Vitamin K to prevent fractures in older women: systematic review and economic evaluation.. *Health Technology Assessment (Winchester, England)* 2009; 13(45):iii-ixi.

360. Singer, B.R., McLauchlan, G.J., Robinson, C.M., Christie, J. Epidemiology of fractures in 15 000 adults. The influence of age and gender. *Journal of Bone & Joint Surgery - British Volume* 1998; 80-B(243):248.

361. Holt, G., Khaw, K.T., Reid, D.M., Compston, J.E., Bhalla, A., Woolf, A.D. et al. Prevalence of osteoporotic bone mineral density at the hip in Britain differs substantially from the US over 50 years of age: implications for clinical densitometry. *British Journal of Radiology* 2002; 75(897):736-742.

362. Stevenson, M., Jones, M.L., De, N.E., Brewer, N., Davis, S., Oakley, J. et al. A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis.. *Health Technology Assessment (Winchester, England)* 2005; 9(22):1-160.

363. Jalava, T., Sarna, S., Pylkkänen, L., Mawer, B., Kanis, J.A., Selby, P. et al. Association between vertebral fracture and increased mortality in osteoporotic patients. *Journal of Bone & Mineral Research* 2003; 18(7):1254-1260.

364. Curtis, L. Unit costs of health & social care. 2011; available from http://www.pssru.ac.uk/archive/pdf/uc/uc2011/uc2011.pdf (accessed May 2012).

365. Torgerson, D.J., Dolan, P. Prescribing by general practitioners after an osteoporotic fracture. *Annals of the Rheumatic Diseases* 1998; 57:378-379.

366. Morrison, L.S., Tobias, J.H. Effect of a case-finding strategy for osteoporosis on bisphosphonate prescribing in primary care. *Osteoporosis International* 2005; 16:71-77.

367. Wilson, D.J. Vertebroplasty for vertebral fracture. On the basis of current evidence, cannot be recommended as first line treatment. *BMJ* 2011; 343:104-105.

368. Finniss, D.G., Kaptchuk, T.J., Miller, F., Benedetti, F. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010; 375(9715):686-695.

369. Benedetti, F., Amanzio, M. The placebo response: how words and rituals change the patient's brain. *Patient Educ Couns* 2011; 84(3):413-419.

370. Benedetti, F., Mayberg, H.S., Wager, T.D., Stohler, C.S., Zubieta, J.K. Neurobiological mechanisms of the placebo effect. *J Neurosci* 2005; 25(45):10390-10402.

371. Amanzio, M., Benedetti, F. Neuropharmacological dissection of placebo analgesia: expectation-activated opioid systems versus conditioning-activated specific subsystems. *J Neurosci* 1999; 19(1):484-494.

372. Levine, J.D., Gordon, N.C., Fields, H.L. The mechanism of placebo analgesia. *Lancet* 1978; 2(8091):654-657.

373. Firanescu, C., Lohle, P.N.M., de Vries, J., Klazen, C.A., Juttmann, J.R., Clark, W. et al. A randomised sham controlled trial of vertebroplasty for painful acute osteoporotic vertebral fractures (VERTOS IV). *Trials* 2011; 12:93.

374. Buchbinder, R., Kallmes, D.F. Vertebroplasty: when randomized placebo-controlled trial results clash with common belief. *Spine Journal* 2010; 10(3):241-243.

375. Harstall, R., Heini, P.F., Mini, R.L., Orler, R. Radiation exposure to the surgeon during fluoroscopically assisted percutaneous vertebroplasty: a prospective study. *Spine* 2005; 30(16):1893-1898.

376. Seibert, J.A. Vertebroplasty and kyphoplasty: Do fluoroscopy operators know about radiation dose, and should they want to know? *Radiology* 2004; 232(3):633-634.

377. Synowitz, M., Kiwit, J. Surgeon's radiation exposure during percutaneous vertebroplasty. *Journal of Neurosurgery Spine* 2006; 4(2):106-109.

378. Kruger, R., Faciszewski, T. Radiation dose reduction to medical staff during vertebroplasty: a review of techniques and methods to mitigate occupational dose. *Spine* 2003; 28(14):1608-1613.

379. Tappero, C., Barbero, S., Costantino, S., Bergui, M., Ropolo, R., Bradac, G. et al. Patient and operator exposure during percutaneous vertebroplasty. *Radiologia Medica* 2009; 114(4):595-607.

380. Cloft, H.J., Easton, D.N., Jensen, M.E., Kallmes, D.F., Dion, J.E. Exposure of medical personnel to methylmethacrylate vapor during percutaneous vertebroplasty. *Ajnr: American Journal of Neuroradiology* 1999; 20(2):352-353.

381. Edidin, A., Ong, K.L., Lau, E., Kurtz, S. Mortality risk for operated and non-operated vertebral fracture patients in the U.S. Medicare population. *Journal of Vascular and Interventional Radiology* 2011; Conference(var.pagings):S7.

382. Lilford, R., Braunholtz, D., Harris, J., Gill, T. Trials in surgery. *Br J Surg* 2004; 91(1):6-16.

383. Appel, N.B., Gilula, L.A. Percutaneous vertebroplasty in patients with spinal canal compromise. *AJR American Journal of Roentgenology* 2004; 182(4):947-951.

384. Baerlocher, M.O., Munk, P.L., Liu, D.M., Tomlinson, G., Badii, M., Kee, S.T. et al. Clinical utility of vertebroplasty: need for better evidence. *Radiology* 2010; 255(3):669-674.

385. Becker, S., Garoscio, M., Meissner, J., Tuschel, A., Ogon, M. Is there an indication for prophylactic balloon kyphoplasty? A pilot study. *Clinical Orthopaedics & Related Research* 2007; 458:83-89.

386. Bian, J., Mudano, A., Allison, J., Briggs, D., Cope, J., Curtis, J. et al. Vertebroplasty/kyphoplasty increases the risk of secondary vertebral compression fractures. *Journal of Bone and Mineral Research* 2006; 21(S1):S105.

387. Boonen, S., Cummings, S., Van Meirhaeghe, J.K., Bastian, L., Tillman, J., Ranstam, J. et al. A randomized trial of balloon kyphoplasty and non surgical care for patients with acute vertebral compression fractures: Two year results. *Osteoporosis International* 2010; 21(Suppl 1):S21-S22.

388. Buchbinder, R., Osborne, R.H., Wark, J.D., Mitchell, P., Wreidt, C., Graves, S. et al. Efficacy and safety of vertebroplasty for treatment of painful osteoporotic vertebral fractures: A randomised double-blind placebo-controlled trial. *Internal Medicine Journal* 2009; 39(Suppl s2):A46.

389. Buchbinder, R., Osborne, R.H., Kallmes, D. Vertebroplasty appears no better than placebo for painful osteoporotic spinal fractures, and has potential to cause harm. *Medical Journal of Australia* 2009; 191(9):476-477.

390. Cummings, S.R., Wardlaw, D., Van, M.J., Bastian, L., Tillman, J.B., Ranstam, J. et al. A randomized trial of balloon kyphoplasty and nonsurgical care for acute vertebral compression fracture. *Bone* 2009; 44(Suppl 1):S51-S52.

391. Figueiredo, N., Barra, F., Moraes, L., Rotta, R., Casulari, L.A. Percutaneous vertebroplasty: a comparison between the procedure using the traditional and the new side-opening cannula for osteoporotic vertebral fracture. *Arquivos De Neuro-Psiquiatria* 2009; 67(2B):377-381.

392. Gray, L.A., Kallmes, D.F. A pilot study of the use of pain questionnaires in vertebroplasty research [Erratum appears in AJNR Am J Neuroradiol. 2009 Oct;30(9):E144]. *Ajnr: American Journal of Neuroradiology* 2009; 30(7):1364-1365.

393. Holden, L., Cheung, G., Chow, E., Finkelstein, J., Danjoux, C. Prospective evaluation of functional status and quality of life in patients undergoing percutaneous vertebroplasty. *International Journal of Cancer* 2002;135.

394. Mao, K.Y., Liu, B.W., Wang, Y., Tao, S., Wang, J.F., Liu, Z.S. et al. Effect of carbonated hydroxyapatite cement for filing vertebral body on the vertebral heights and pain in patients with osteoporotic vertebral compression fractures. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2007; 11(1):188-190.

395. Ramaswamy, D., Teitelbaum, G.P., Horwitz, D.A., Ehresmann, G.R. Comparison of efficacy of percutaneous vertebroplasty with polymethylmethacrylate in providing pain relief in patients with acute vs chronic osteoporotic vertebral compression fractures. *Arthritis and Rheumatism* 2000; 43(9):S197.

396. Smith, F.W., Boonen, S., Van, M.J., Bastian, L., Wardlaw, D. A randomized trial of balloon kyphoplasty and nonsurgical care for acute vertebral compression fracture: 2-year results. *CardioVascular and Interventional Radiology* 2009; 32(Suppl 2):313-314.

397. ClinicalTrials.gov. Quality of life after vertebroplasty versus conservative treatment in patients with painful osteoporotic fracture. *Internet* 2012; Available from http://clinicaltrials.gov/ct2/show/NCT00994032 (accessed Jan. 2012).

398. ClinicalTrials.gov. FREE study - Fracture Reduction Evaluation. *Internet* 2012; Available from http://clinicaltrials.gov/ct2/show/study/NCT00211211?term=00211211&rank=1 (accessed Feb. 2012).

399. Buchbinder, R., Osborne, R.H., Ebeling, P.R., Wark, J.D., Mitchell, P., Wriedt, C. et al. Supplement to: Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic fractures. N Engl J Med 2009;361:557-48. *Internet* 2009; Available from http://www.nejm.org/action/showSupplements?doi=10.1056%2FNEJMoa0900429&viewType=Popup&viewClass=Suppl (accessed Mar. 2012).

400. Wardlaw, D., Boonen, S., Bastian, L., Van Meirhaeghe, J., St Jan, A.Z. An international multicenter randomized comparison of balloon kyphoplasty and nonsurgical care in patients with acute vertebral body compression fractures. *Journal of Bone and Mineral Research* 2007; 22(7):1119.