NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL (STA)

Romosozumab for treating severe osteoporosis [ID3936]

Appraisal Committee Meeting – 4 November 2021
1st Committee meeting

The **final scope** and **final stakeholder list** are available on the NICE website.

Pre-technical engagement documents

1. **Company submission summary** from UCB Pharma

2. **Clarification questions and company responses**
   a. Clarification questions
   b. Clarification responses

3. **Patient group, professional group and NHS organisation submissions** from:
   a. British Society for Rheumatology
   b. Royal Osteoporosis Society

4. **Evidence Review Group report** prepared by **Kleijnen Systematic Reviews Ltd**

5. **Evidence Review Group report – factual accuracy check**

6. **Evidence Review Group response to FAC**

7. **Appraisal Committee Meeting presentation slides** – to follow

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Romosozumab for treating postmenopausal severe osteoporosis [ID3936]

Document A

Company evidence submission summary for committee

UCB Pharma Ltd confirm that all information in the submission summary is an accurate summary or replication of evidence in the main submission and accompanying appendices and that wherever possible a cross reference to the original source is provided.

June 2021

File name	Version	Contains confidential information	Date
Document A	Final	Yes	25th June 2021

Instructions for companies

This is the template you should use to summarise your evidence submission to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. This document will provide the appraisal committee with an overview of the important aspects of your submission for decision-making.

This submission summary must not be longer than 25 pages, excluding the pages covered by this template. If it is too long it will not be accepted. Please submit a draft summary with your main evidence submission. The NICE technical team may request changes later.

When cross referring to evidence in the main submission or appendices, please use the following format: Document, heading, subheading (page X).

For all figures and tables in this summary that have been replicated, cross refer to the evidence from the main submission or appendices in the caption in the following format: Table/figure name – document, heading, subheading (page X).

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

Highlighting in the template (excluding the contents list)

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables and figures

Submission summary

A.1 Health condition

Osteoporosis is a highly prevalent disease, with an estimated 3.5 million individuals over the age of 50 in the UK living with osteoporosis.[1-3] Osteoporosis is characterised by low bone mass and deterioration in bone microarchitecture, resulting from an imbalance between bone resorption and formation in the naturally occurring bone remodelling cycle.[2] This imbalance increases with age and is most common among postmenopausal women.[2]

Women with osteoporosis are at increased risk of fragility fractures – fractures occurring from low trauma, due to reduced bone strength.[4-7] One third of postmenopausal women will suffer a fragility fracture due to osteoporosis in their lifetime.[1, 8] A fracture is a major risk factor for future fractures. The relative risk of a future fracture sharply increases and is highest in the two years following a fracture, during which time women are at *imminent risk* of another fracture.[9-11] A postmenopausal woman who has recently suffered a major osteoporotic fracture (MOF; fracture of the hip, spine, wrist, or humerus) is over five times more likely to suffer another fracture within one year.[12-14]

Fragility fractures result in considerable disability and pain for patients, as well as significant impairments in mobility, reduced independence and increased frailty.[15-19] Fragility fractures are also associated with significantly increased mortality.[20-23] Despite existing treatments there is a major unmet need for an effective, fast-acting and easy to use treatment to be made available to patients immediately following a recent MOF, to interrupt and prevent the cycle of further fragility fractures, and their associated morbidity and mortality.[10, 13, 24]

A.2 Clinical pathway of care

The current first-line pharmaceutical treatments for women with osteoporosis and at high risk of a fragility fracture are oral bisphosphonates, such as once weekly (QW) alendronate or risedronate, which act by decreasing bone resorption.[25] Bisphosphonates are anti-resorptive treatments, which inhibit osteoclast activity, reducing the breakdown of bone via the bone remodelling process.[26-31] While these treatments do not have a direct impact on bone formation, bone formation is reduced secondary to the reduction in bone resorption.

Patients at higher risk of fracture, or those who are unable to tolerate oral bisphosphonates, may instead be treated with intravenous (IV) bisphosphonates, such as IV zoledronate once yearly.[25, 32] Denosumab (60 mg subcutaneous [SC] injection once every six months) and raloxifene (60 mg tablet once daily [QD]) are alternative anti-resorptive treatments for women unable to tolerate or who have a poor response to bisphosphonates.[32, 33] These therapies also work via reductions in bone resorption.[34]

Teriparatide (20 micrograms [µg], SC QD injection for 24 months), a bone-forming agent, is used as an alternative secondary prevention treatment. for a subset of women who are unable to tolerate, are contraindicated for or have a poor response to alendronate and risedronate. Women must also be 65 years or older and have a T-score of ≤–4.0, or a T-score of ≤–3.5 plus more than two fractures, or are aged 55–64 years and have a T-score of ≤–4 plus more than two fractures.[35] By this stage, patients who receive treatment with teriparatide will have suffered multiple fractures, experienced the associated pain, disability, increased frailty and reduced health-related quality of life (HRQoL) associated with fractures and are at significantly increased

risk of mortality.[15-20, 36-42] The use of teriparatide is limited to 24 months once in a lifetime, due to preclinical studies highlighting concerns that longer treatment periods may increase the risk of developing osteosarcoma.[43-45]

A summary of the NICE guidelines for osteoporosis treatments can be found in Table 1.

Table 1: Summary of NICE guidelines and guidance for osteoporosis

<table>
<tr><th>NICE guidance</th><th>Summary of NICE recommendation</th></tr>
<tr><td>TA161 (2008; reissued 2018).[35]</td><td>Strontium ranelate [now withdrawn] and raloxifene as alternative treatment options for the secondary prevention of osteoporotic fractures in postmenopausal women who:
Are unable to comply with the administration of alendronate and risedronate, or are intolerant/contraindicated to these
Have a combination of T-score, age and number of independent clinical risk fractures[a] as indicated in the table below:
<table>
<tr><th rowspan="2">Age (years)</th><th colspan="3">Number of independent clinical risk factors for fracture</th></tr>
<tr><th>0</th><th>1</th><th>2</th></tr>
<tr><td>50–54</td><td>NA[a]</td><td>−3.5</td><td>−3.5</td></tr>
<tr><td>55–59</td><td>−4.0</td><td>−3.5</td><td>−3.5</td></tr>
<tr><td>60–64</td><td>−4.0</td><td>−3.5</td><td>−3.5</td></tr>
<tr><td>65–69</td><td>−4.0</td><td>−3.5</td><td>−3.0</td></tr>
<tr><td>70–74</td><td>−3.0</td><td>−3.0</td><td>−2.5</td></tr>
<tr><td>75 or older</td><td>−3.0</td><td>−2.5</td><td>−2.5</td></tr>
<tr><td colspan="4">[a]Treatment with raloxifene or strontium ranelate is not recommended</td></tr>
</table>
Teriparatide as an alternative for secondary prevention of osteoporotic fragility fracture in postmenopausal women who:
Are unable to take alendronate and risedronate, or are intolerant/contraindicated to alendronate and risedronate or have had an unsatisfactory response to alendronate or risedronate and
Are 65 years or older and have a T-score of ≤–4.0, or a T-score of ≤–3.5 plus more than two fractures, or are aged 55–64 years and have a T-score of ≤–4 plus more than two fractures</td></tr>
<tr><td>TA204 (2010)[33]</td><td>Denosumab for the primary prevention of osteoporotic fragility fractures in postmenopausal women with increased risk of fracture who:
Are unable to comply with the administration of alendronate and either risedronate/etidronate, or are intolerant/contraindicated to these and
Have a combination of T-score, age and number of independent clinical risk fractures[a] as indicated in the table below:
<table>
<tr><th rowspan="2">Age (years)</th><th colspan="3">Number of independent clinical risk factors for fracture</th></tr>
<tr><th>0</th><th>1</th><th>2</th></tr>
<tr><td>65-69</td><td>a</td><td>−4.5</td><td>−4.0</td></tr>
<tr><td>70-74</td><td>−4.5</td><td>−4.0</td><td>−3.5</td></tr>
<tr><td>75 or older</td><td>−4.0</td><td>−4.0</td><td>−3.0</td></tr>
<tr><td colspan="4">[a]Treatment with denosumab is not recommended</td></tr>
</table></td></tr>
</table>

	Denosumab for the secondary prevention of osteoporotic fragility fractures for postmenopausal women with increased risk of fracture who: • Cannot comply with the administration of alendronate and either risedronate/etidronate, or are intolerant/contraindicated to these treatments
TA464 (2017; reissued 2019)[25]	Oral bisphosphonates (alendronic acid, ibandronic acid and risedronate sodium) and intravenous bisphosphonates (ibandronic acid and zoledronic acid) are recommended, within their marketing authorisations, as options for treating osteoporosis in adults: • who are eligible for risk assessment as defined in NICE's guideline on osteoporosis (CG146; recommendations 1.1 and 1.2) and NICE's quality standard on osteoporosis (QS149) and • who have been assessed as being at higher risk of osteoporotic fragility fracture using the methods recommended in NICE's guideline on osteoporosis (CG146; recommendations 1.3 to 1.12) and NICE's quality standard on osteoporosis (QS149) and • when bisphosphonate treatment is appropriate, taking into account their risk of fracture, their risk of adverse effects from bisphosphonates, and their clinical circumstances and preferences. The choice of treatment should be made on an individual basis after discussion between the responsible clinician and the patient, or their carers, about the advantages and disadvantages of the treatments available. If generic products are available, start treatment with the least expensive formulation, taking into account administration costs, the dose needed and the cost per dose.

Abbreviations: CG: Clinical Guideline; NICE: National Institute for Health and Care Excellence; QS: Quality Standard; TA: technology appraisal.

A.3 Equality considerations

Romosozumab is only licensed for use in postmenopausal women, not men; however, osteoporosis is four times more likely to occur in women than men.[46] Fragility fractures do not affect all patients equally. Social deprivation is predictive of increased fracture risk, higher mortality in the year following a hip fracture, and among survivors, longer hospital stays and risk of re-admission.[47-50] One study of 218,907 admissions with an index hip fracture (mean age 82.8 years; 72.6% female) found that patients in the most deprived quintile in England experienced a 24% increase in mortality (age-sex-comorbidity-adjusted odds ratio (OR):1.24 [1.20, 1.28], p<0.001; Q5 versus Q1) one year following a hip fracture, compared to patients in the least deprived quintile (measured using the Index of Multiple Deprivation quintiles).[48] Between 2001 and 2015, the health equality gap for hip fracture incidence marginally widened among women.[47]

A.4 The technology

Table 2 Technology being appraised – B.1.2 (Page 13)

UK approved name and brand name	Romosozumab (EVENITY®)
Mechanism of action	Romosozumab is a monoclonal antibody that binds to and inhibits sclerostin.[51] Inhibition of sclerostin has a dual effect on bone. It stimulates bone formation through promoting increased osteoblast number and activity, as inhibition of sclerostin activates Wnt signalling, and also reduces bone resorption through changing the expression of osteoclast mediators.[51]
Marketing authorisation/CE mark status	Following an application to the European Medicines Agency (EMA) under the centralised procedure,

	marketing authorisation was granted on 9th December 2019.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Romosozumab is indicated for the treatment of severe osteoporosis in postmenopausal women at high risk of fracture.[52] Romosozumab is contraindicated for patients with:[52] • Hypersensitivity to the active substance(s) or to any of the excipients • Hypocalcaemia • History of myocardial infarction or stroke
Method of administration and dosage	Romosozumab is administered as two subcutaneous injections at a total dose of 210 mg once monthly for a 12-month course of treatment. Following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months.
Additional tests or investigations	NA
List price and average cost of a course of treatment	List price of romosozumab: £427.75 for each monthly dose consisting of two pre-filled pens. Cost for a fixed-duration 12-month treatment (based on list price): £5,133.
Patient access scheme (if applicable)	A patient access scheme (PAS) has been proposed for romosozumab. The proposed romosozumab with-PAS net price is £[redacted] per monthly dose, equivalent to a percentage discount of [redacted]%.

Abbreviations: EMA: European Medicines Agency; mg: milligram; PAS: patient access scheme; SmPC: summary of product characteristics; UK: United Kingdom.

A.5 Decision problem and NICE reference case

This submission focuses on a population that is part of the marketing authorisation of romosozumab. Romosozumab is indicated for the treatment of severe osteoporosis in postmenopausal women at high risk of fracture.[52] Within the license for romosozumab, the target patient population considered in this submission is patients who have:

- Experienced a recent MOF within the past 24 months; and
- Thus, are at imminent risk of another fragility fracture.

The proposed patient population focusses on women with the greatest unmet need, and for whom romosozumab is expected to provide substantial (or pronounced) clinical benefit. The decision problem considered within this submission is detailed in Table 3.

Table 3. The decision problem – B.1.1 (page 11–12)

	Final scope issued by NICE/reference case	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	• Postmenopausal women with severe osteoporosis at high risk of fracture	• Postmenopausal women with severe osteoporosis who are at high risk of fracture and who have: o Experienced a recent MOF within 24 months; and o Thus, are at imminent risk of another fragility fracture	• The submission positions romosozumab for use in the subgroup of the licenced population who have greatest unmet need, and for whom romosozumab is expected to provide substantial clinical benefit
Intervention	• Romosozumab	• Romosozumab for 12 months, followed by alendronate	• Romosozumab is licensed as a 12-month course of treatment • The SmPC for romosozumab states that "following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months"
Comparator(s)	• Bisphosphonates (alendronic acid, risedronate sodium, ibandronic acid and zoledronic acid) • Non-bisphosphonates including antiresorptive agents (denosumab and raloxifene) and anabolic agents (teriparatide) • No active treatment	• The base case comparison is with alendronate, using the head-to-head ARCH study • Scenario analyses are provided against all other comparators listed in the scope, using the NMA, except ibandronic acid	• No trials of the licensed dose of ibandronate found to be included in the NMA, therefore comparisons could not be conducted
Outcomes	• Osteoporotic fragility fracture • Bone mineral density • Mortality • Adverse effects of treatment • Health-related quality of life	• In line with the final NICE scope	• In line with the final NICE scope

Abbreviations: MOF: major osteoporotic fracture; NICE: National Institute for Health and Care Excellence; NMA: network meta-analysis; SmPC: summary of product characteristics; UK: United Kingdom.

A.6 Clinical effectiveness evidence

The clinical effectiveness evidence for romosozumab in severe osteoporosis in postmenopausal women is provided from three Phase III clinical trials: ARCH, FRAME and STRUCTURE.

The clinical evidence presented in support of this submission is principally provided by ARCH, a pivotal study that provides direct comparative evidence for romosozumab/alendronate versus alendronate alone.[53] This superiority study against the widely-used comparator, alendronate, was designed to show evidence of fracture risk reduction along with superior bone mineral density (BMD) outcomes. ARCH was a Phase III, multicentre, randomised, double-blind, alendronate-controlled trial in postmenopausal women with severe osteoporosis and a prior fragility fracture. Patients received either the once monthly (QM) SC dose of romosozumab 210 mg or oral alendronate 70 mg QW for 12 months, followed by alendronate 70 mg QW in both treatment arms.

The ARCH trial provides evidence for romosozumab in its expected position in the clinical pathway: a first-line therapy in patients who have previously suffered a MOF. Efficacy outcomes reported in ARCH are clinically relevant and include incidence of clinical, vertebral, non-vertebral and hip fracture and percentage change from baseline in BMD. As such, data from ARCH were used as the principal clinical effectiveness evidence for the economic modelling in this submission.

FRAME and STRUCTURE were considered as supportive clinical evidence in this submission because FRAME did not include a patient population aligned to where romosozumab is expected to be used in NHS clinical practice and STRUCTURE was not designed to evaluate fracture outcomes.[54, 55]

Table 4: Clinical effectiveness evidence

Study title	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
Study design	• International, multicentre, randomised, double-blind, active-controlled, parallel-group, Phase III	• International, multicentre, randomised, double-blind, placebo-controlled, parallel-group, Phase III	• International, multicentre, randomised, open-label, active-controlled, parallel-group, Phase III
Population	• Ambulatory postmenopausal women with osteoporosis • Aged 55–90 years • Prior fragility fracture	• Postmenopausal women with osteoporosis • Aged 55–90 years	• Postmenopausal women with osteoporosis transitioning from 3 years of bisphosphonate therapy • Aged 55–90 years • Prior fragility fracture
Intervention(s)	• Romosozumab (210 mg) QM SC for 12 months followed by open-label oral alendronate (70 mg) QW for at least 12	• Romosozumab (210 mg) QM SC for 12 months followed by open-label denosumab (60 mg) SC Q6M for 24 months (until study end)	• Romosozumab (210 mg) QM SC for 12 months

Study title	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
	months (until study end)		
Comparator(s)	• Oral alendronate (70 mg) QW for 12 months followed by open-label alendronate (70 mg) for at least 12 months (until study end)	• Placebo QM SC for 12 months followed by open-label denosumab (60 mg) Q6M SC for 24 months (until study end)	• Daily SC teriparatide (20 µg) for 12 months
Outcomes specified in the decision problem	• Cumulative incidence of new vertebral fracture • Cumulative incidence of clinical fracture • Incidence of fractures (non-vertebral, all fractures, new or worsening vertebral, major non-vertebral, hip, MOF) • Percent change in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs	• Incidence of a new vertebral fracture • Cumulative incidence of non-vertebral fracture, major non-vertebral fracture, clinical fracture, hip fracture, new or worsening vertebral fracture, MOF and multiple new or worsening vertebral fractures • Percent change from baseline in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs	• Percent change from baseline in BMD at LS, TH, and FN • Finite element analysis of the hip[a] • AEs
Reference to section in submission	• Section B.2	• Section B.2	• Section B.2

Footnotes: [a] Finite element analysis of the hip results are available in Appendix L.6.
Abbreviations: AE: adverse event; BMD: bone mineral density; BPI: Brief Pain Inventory; EQ-5D-5L: EuroQoL-5 Dimensions-5 Levels Health Survey; FN: femoral neck; LAD: Limited Activity Days; LS: lumbar spine; mg: milligram; MOF: major osteoporotic fracture; OPAQ-SV: Osteoporosis Assessment Questionnaire Short Version; PRO: patient report outcome; QM: once monthly; Q6M: once every six months; QW: once weekly; SC: subcutaneous; TH: total hip; µg: microgram.
Sources: ARCH Clinical Study Report[56]; FRAME Clinical Study Report[57]; STRUCTURE Clinical Study Report[58]

A.7 Key results of the clinical effectiveness evidence

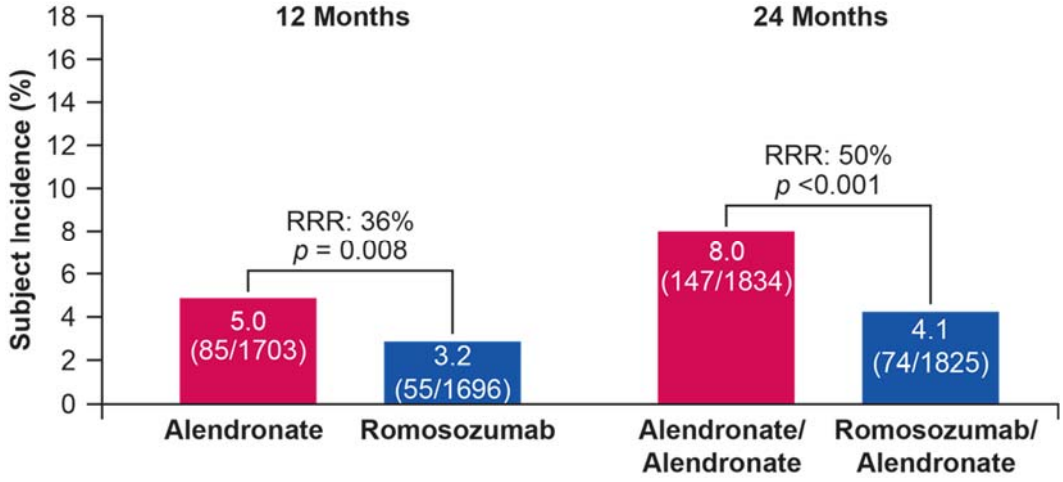
A.7.1 ARCH primary endpoints: cumulative incidence of new vertebral fracture at 24 months and cumulative incidence of clinical fracture at primary analysis

Romosozumab significantly reduced the incidence of new vertebral fractures at 24 months versus alendronate, which was already established at Month 12

ARCH was designed as an event-driven trial. The primary analysis for ARCH was performed after all patients had completed their Month 24 visit and at least 330 patients had confirmed events of clinical fracture (composite of non-vertebral fracture and clinical vertebral fracture [a suspected vertebral fracture that is brought to medical attention and confirmed]).

Patients in the romosozumab arm had a 36% lower relative risk of vertebral fractures compared to alendronate by Month 12 (nominal p=0.008), demonstrating the rapid effect of romosozumab at reducing fracture risk (Figure 1). The absolute risk reduction [redacted][56] This reduction in new vertebral fracture risk versus alendronate was sustained and increased through Month 24 in the romosozumab/alendronate arm (RRR: 50%; [redacted] adjusted and nominal p<0.001), meeting the co-primary endpoint for ARCH.[56]

Figure 1: Incidence of new vertebral fracture at 12 and 24 months in ARCH[a]



Footnotes: [a] Number of patients in each arm is the number of subjects in the primary analysis set for vertebral fractures.
Abbreviations: RRR: relative risk reduction.
Source: Adapted from ARCH clinical study report.[56]

The proportion of patients experiencing a clinical fracture by the time of primary analysis (which occurred at a median follow-up of 33 months) was significantly lower in the romosozumab/alendronate arm

Romosozumab treatment showed a rapid effect in reducing the risk of fracture, with the risk of clinical fracture significantly lower in patients treated with romosozumab compared to alendronate at Month 12 (Figure 2).

As demonstrated in Figure 3 there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first clinical fracture by Month 12. At the time of primary analysis, patients treated with romosozumab/alendronate had a lower cumulative incidence of clinical fracture (9.7%) compared to the alendronate/alendronate group (13.0%; nominal and adjusted p<0.001) (Figure 2).[53, 56] This equated to a 27% lower relative risk of clinical fracture in the romosozumab/alendronate group than alendronate alone, meeting the co-primary endpoint for ARCH.

Figure 2: Incidence of clinical fracture at 12 and 24 months, and primary analysis in ARCH

Abbreviations: RRR: relative risk reduction.
Source: Adapted from ARCH clinical study report.[56]

Figure 3: Kaplan-Meier curves for time to first clinical fracture[a]

Footnotes: [a] Risks presented are based on an LOCF method for patients with missing fracture status. For Kaplan-Meier curves in the time-to-event analysis, data from patients who withdrew or reached the end of the reporting period without having a fracture were carried forward from the last observation time.
Abbreviations: LOCF: last observation carried forward; N: number of subjects randomised; n: number of subjects at risk for event at time point of interest.
Source: Adapted from ARCH clinical study report.[56]

A.7.2 ARCH secondary endpoints: fracture results

Cumulative incidence of non-vertebral, MOF and hip was reduced in the romosozumab/alendronate group compared to alendronate alone at primary analysis (Figure 4). Patients initially treated with romosozumab showed 19% lower risk of non-vertebral fracture compared to alendronate, with an incidence of fracture of 8.7% compared to 10.6% (adjusted p=0.040) at primary analysis.[56] Incidence of MOF (including fracture of the hip, forearm and humerus that are not associated with a pathologic fracture regardless of trauma severity, and clinical vertebral fractures) was [redacted] in the romosozumab/alendronate group versus [redacted] in the alendronate group (RRR: 32%; nominal p<0.001) at primary analysis.[56] Numerical, non-significant, reductions of similar magnitude were already present by Month 12 (non-significance at this earlier timepoint is expected from the event-driven nature of the study design). Incidence of hip fracture was 2.0% in the romosozumab/alendronate group versus 3.2% in the alendronate group (RRR: 38%; nominal p=0.015) at primary analysis.

Figure 4: Incidence of non-vertebral, major osteoporotic and hip fractures at primary analysis

Footnotes: [a] Adjusted 2-sided p value presented for incidence of non-vertebral fractures
Abbreviations: RRR, relative risk reduction.
Source: Adapted from ARCH clinical study report.[56]

A.7.3 Bone mineral density results in ARCH

Treatment with romosozumab resulted in rapid and significant improvement in BMD at the lumbar spine, total hip and femoral neck at Month 12 (Figure 5) compared to alendronate. Following transition from romosozumab to alendronate treatment, this improvement was maintained through Month 36. In a subgroup of patients in which BMD was assessed every 6 months, the improvement in BMD in response to romosozumab was evident by Month 6 (the earliest time of assessment; presented in Figure 11, Section B.2.6.3 of Document B) of treatment (adjusted p<0.001 for all comparisons), indicative of the rapid onset of treatment effect with romosozumab.[53] As can be noted from the data, romosozumab achieved higher BMD gains at Month 12 than alendronate achieved at Month 36.

Figure 5: Mean change from baseline in BMD through Month 36 in ARCH

Footnotes: Data are least square mean percentage changes in BMD based on LOCF. *p< for all comparisons.
Abbreviations: BMD: bone mineral density; LOCF: last observation carried forward.
Source: Adapted from ARCH clinical study report[56]

A.7.4 Summary of primary clinical efficacy results from FRAME

By Month 12, FRAME demonstrated statistically significant reductions in new vertebral fractures for romosozumab compared with placebo (RRR: 73%; [redacted] adjusted p<0.001). Similarly, patients in the romosozumab/denosumab arm showed a significant 75% reduction in relative risk of new vertebral fracture compared to the placebo/denosumab arm ([redacted] incidence of new vertebral fracture: 0.6% versus 2.5%; 95% CI: 60, 84; adjusted p<0.001) at Month 24.[54] Romosozumab also reduced the risk of clinical fracture (non-vertebral and clinical vertebral fracture) by 36% compared with placebo through Month 12 (adjusted and nominal p=0.008) and to 33% through Month 24 (adjusted p=0.096, nominal p=0.002).[54]

Romosozumab/denosumab also numerically reduced major non-vertebral, new or worsening vertebral and other fractures through Month 24 compared to placebo/denosumab, although these were not considered statistically significant due to the endpoint testing sequence.

A.7.5 Summary of primary clinical efficacy results from STRUCTURE

STRUCTURE provides BMD and estimated bone strength data comparing romosozumab and teriparatide in a population with severe osteoporosis and who received an oral bisphosphonate before transitioning to the bone-forming agent. In STRUCTURE, the mean percentage change from baseline up to Month 12 in BMD at the total hip was 3.2% higher (95% CI: 2.7, 3.8; adjusted p<0.0001) in the romosozumab group (2.6% [95% CI: 2.2, 3.0]) compared to teriparatide (–0.6% [95% CI: –1.0, –0.2]). Superior gains in BMD with romosozumab compared to teriparatide were also observed at the lumbar spine and femoral neck.[55]

A.7.6 Health-related quality of life

HRQoL data were available from ARCH and FRAME.[56, 57] In both studies, [redacted] were observed between the treatment groups.[56, 57] This was expected because the HRQoL data were collected at predetermined, discrete time points irrespective of fracture occurrence during the trial and always related to one of the investigated treatments. It is also important to note that the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited. [redacted] a decline in HRQoL was observed following a fracture on both treatments. By preventing fragility fractures, romosozumab is expected to prevent future HRQoL decrements resulting from a fracture.

A.8 Evidence synthesis

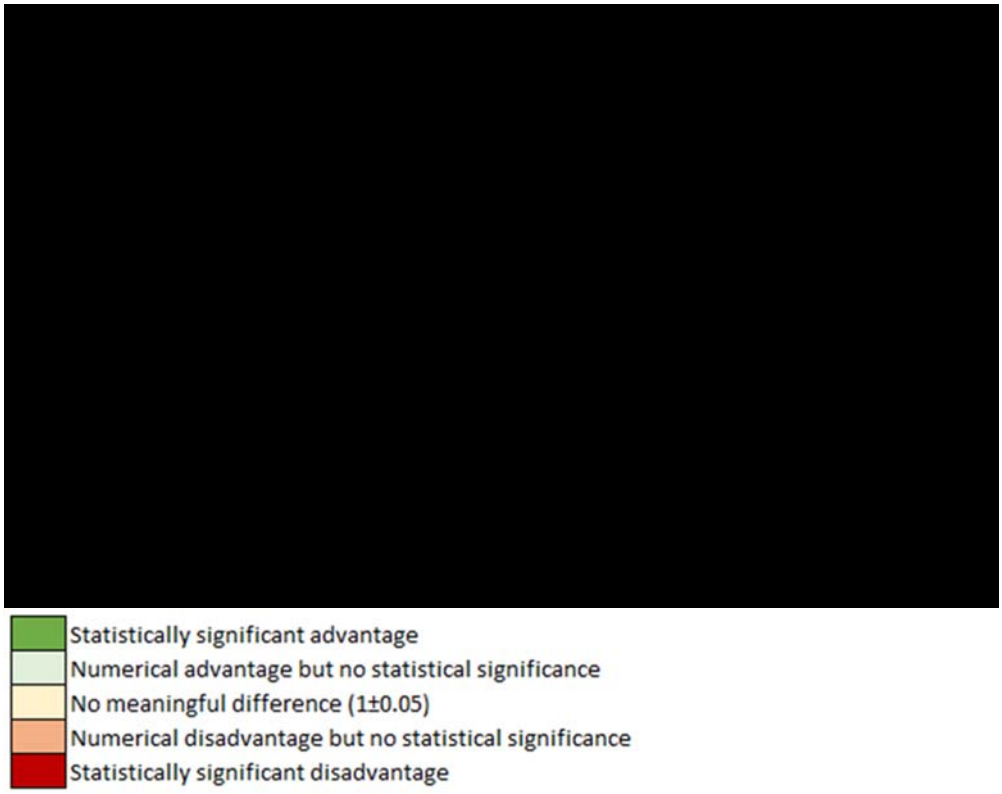
A network meta-analysis (NMA) was conducted to compare the efficacy of romosozumab/alendronate versus other bisphosphonates (alendronate, risedronate, ibandronate, zoledronate), teriparatide, denosumab, and raloxifene; the NMA was not undertaken specifically for this appraisal and additionally includes abaloparatide, which is not licensed in the UK and is not a comparator in the NICE final scope for this appraisal. It was intended for the NMA to include ibandronate, however ibandronate could not be joined to any of the networks for fracture outcomes, as no trials provided evidence for the licensed dose regimen.

The NMA used available data from randomised controlled trials (RCTs) identified in the systematic literature review (SLR), which included placebo-controlled and head-to-head RCTs with at least 12 months follow-up investigating the treatment of postmenopausal women with

osteoporosis at increased risk of fracture. It should be noted that the results from the NMA are confounded by the heterogeneity of the included trial populations and differences in trial designs between ARCH and the other trials included.

Figure 6 below describes the significance of romosozumab versus the comparators per time-point specific fracture endpoints. The results demonstrate that romosozumab was significantly more effective than or at least as good as most of the comparators included in the NMA. Romosozumab for 12 months followed by alendronate (romosozumab/alendronate) was the treatment [redacted] reported as the treatment with the [redacted] or [redacted] probability of being the most effective treatment.

Figure 6: Relative risk of romosozumab versus comparators for time-point specific fracture endpoints



Abbreviations: FE: fixed effects; Fx: fractures; ITT: intention-to-treat; mo: months; NA: not applicable; ROMO: romosozumab; VFx: vertebral fractures;.

New vertebral fracture

The fixed effects model demonstrated that romosozumab significantly reduced the risk of new vertebral fractures at 12 months compared to placebo (RR [redacted]), raloxifene (RR and alendronate (RR [redacted]). In the random effects [redacted]) model, romosozumab was significantly better compared to placebo (RR [redacted]).

In the fixed effects model, romosozumab/alendronate significantly reduced the risk of new vertebral fractures at 24 months compared to placebo (RR [redacted]), raloxifene (RR [redacted]), alendronate (RR [redacted]), risedronate (RR [redacted]) and zoledronate (RR [redacted]). In the random effects model the results are comparable to the fixed effects model with romosozumab/alendronate being significantly better compared to placebo (RR [redacted]), raloxifene (RR [redacted]), alendronate (RR [redacted]) and risedronate (RR [redacted]).

In the fixed effects model, romosozumab/alendronate significantly reduced the risk of new vertebral fractures at 36 months compared to placebo (RR [redacted]), raloxifene (RR [redacted]), alendronate (RR [redacted]) and risedronate (RR [redacted]). In the random effects model the results are comparable to the fixed effects model with romosozumab/alendronate being significantly more effective than placebo (RR [redacted]), raloxifene (RR [redacted]), alendronate (RR [redacted]) and risedronate ([redacted]).

Non-vertebral fracture

In the fixed effects model of non-vertebral fractures at 12 months, romosozumab significantly reduced the risk of non-vertebral fractures compared to placebo (RR [redacted]) and raloxifene (RR [redacted]). In the results for the random effects model, romosozumab showed no statistically significant differences but showed a trend of reduced fracture risk compared to all treatments.

In the fixed effects model at 24 months, romosozumab/alendronate significantly reduced the risk of non-vertebral fractures compared to placebo (RR [redacted]), raloxifene (RR [redacted]), zoledronate (RR [redacted]) and denosumab (RR [redacted]). In the results for the random effects model, romosozumab/alendronate showed no statistically significant differences but showed a trend of reduced fracture risk against all comparators.

In the fixed effects model at 36 months, romosozumab/alendronate significantly reduced the risk of non-vertebral fractures compared to placebo (RR [redacted]), raloxifene (RR [redacted]), alendronate (RR [redacted]), zoledronate (RR [redacted]) and denosumab (RR [redacted]). In the results for the random effects model, romosozumab/alendronate significantly reduced the risk of non-vertebral fractures compared to placebo (RR [redacted]).

Hip fracture

No study was powered to detect significant improvements in hip fracture risk and most studies in the network had a limited number of hip fractures. In the fixed effects model for the ITT population for hip fractures at 12 months, romosozumab showed a trend of reduced hip fracture risk without reaching statistical significance, except against denosumab. These outcomes hold in the random effects model analysis. It should be noted that not many studies were included in this analysis and that the event rate was low in all included trials, which impacts the results for this endpoint.

At 24 months the results of the fixed effects model showed that romosozumab/alendronate significantly reduced the risk of hip fractures compared to placebo (RR [redacted]). The results for the random effects model showed that romosozumab/alendronate was numerically better compared to placebo, raloxifene, alendronate, zoledronate and abaloparatide. Note that the event rate for teriparatide and abaloparatide was low in the respective RCTs, which may affect the results for this endpoint.

At 36 months the results of the fixed effects model showed that romosozumab/alendronate significantly reduced the risk of hip fractures for placebo (RR [redacted]), raloxifene (RR [redacted]), and alendronate (RR [redacted]). The results for the random effects model showed that romosozumab/alendronate significantly reduced the risk of hip fractures compared to placebo (RR [redacted]).

Conclusion

Across the NMA, romosozumab or romosozumab/alendronate was the treatment [redacted] reported as the treatment with the [redacted] or [redacted] probability of being the most effective treatment across all fracture sites and timepoints considered.

A.9 Key clinical issues

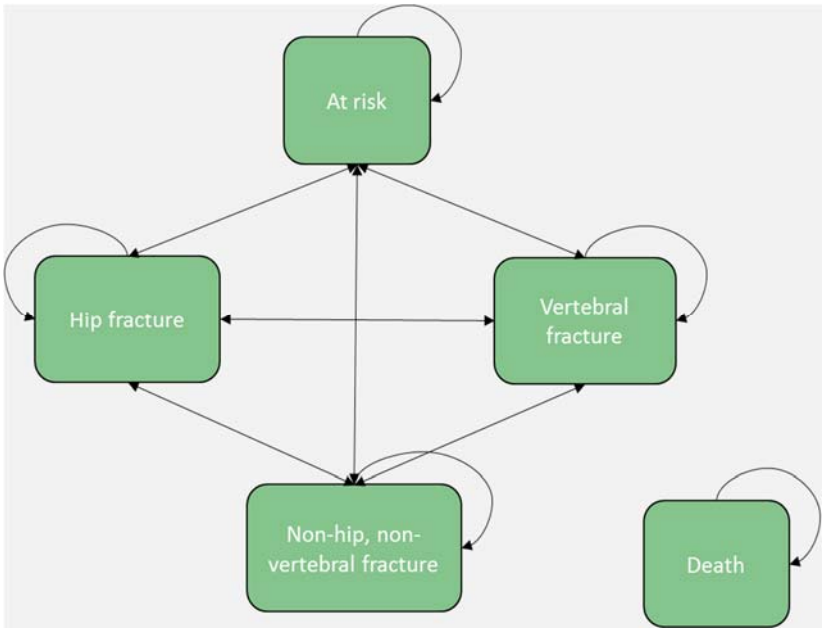
- Whilst ARCH provided a strong clinical evidence base for the use of romosozumab versus alendronate in postmenopausal women with a previous fragility fracture, it does not provide direct comparative evidence for romosozumab versus other therapies for osteoporosis, such as other bisphosphonates (risedronate, zoledronate and ibandronate), raloxifene, denosumab or teriparatide. However, bisphosphonates are considered to be of similar efficacy within the class, as discussed in a recent NICE appraisal of bisphosphonates (NICE TA464), and it is therefore reasonable to assume that alendronate is representative of the bisphosphonate class.[25]
- The results from the NMA are limited by the quantity and quality of the data available from the trials. The main issues of concern relate to differences in the definition of fracture outcomes and how they were measured during the trial. Some trials recorded fracture outcomes as the number of patients with a fracture and others as the time to fracture. Moreover, there were differences in patient populations (as some trials included patients with a prevalent osteoporotic fracture and others did not). BMD, T-scores, treatment history and received concomitant medications at baseline also varied across included RCTs. Follow-up periods also differed and ranged from 12 to 72 months, but approximately 50% had a follow-up period of 36 months. The two romosozumab trials included treatment sequencing where all patients changed to a different treatment: ARCH switched to alendronate at Month 12 and FRAME switched to denosumab at Month 12.
- In both ARCH and FRAME studies, [redacted] were observed for HRQoL data between the treatment groups.[56, 57] This was expected because the HRQoL data were collected at predetermined, discrete time points irrespective of fracture occurrence during the trial. It is also important to note that the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited. [redacted].[53, 54] By preventing fragility fractures, romosozumab is expected to prevent future HRQoL decrements resulting from a fracture.

A.10 Overview of the economic analysis

A *de novo* individual patient simulation state transition model was developed.[59, 60] This approach was validated by experts and deemed acceptable and consistent with models developed previously for anti-osteoporotic drugs when an early version of the model was independently reviewed under the NICE Preliminary Independent Model Advice (PRIMA) process.[61, 62] The model adheres to the recommendations on modelling in osteoporosis by European Society for Clinical and Economic Aspects of Osteoporosis (ESCEO) and International Osteoporosis Foundation (IOF).[63] Furthermore, the cost-effectiveness model has been published in two peer-reviewed manuscripts and formed part of the evaluation that led to reimbursement of romosozumab in Sweden (Tandvårds- och läkemedelsförmånsverket, TLV, The Swedish Dental and Pharmaceutical Benefits Agency) and Scotland (Scottish Medicines Consortium, SMC); evaluations in other EU countries are ongoing.[59, 64-66]

The model consisted of five health states: “at risk” of fracture, hip fracture, vertebral fracture, other osteoporotic fracture (non-hip, non-vertebral; NHNV), and death (Figure 7). The fracture sites included are those that are most strongly associated with osteoporosis and these are the fracture sites included in commonly used risk assessment tools, including FRAX® (Fracture Risk Assessment Tool).[67] The model did not restrict the sequence or number of fractures experienced, reflecting the nature of the disease. All patients started the model in the “at risk” health state. At the end of each cycle patients either moved into the one of the fracture states, remained in their current health state without new fracture, or died. If a patient died, they moved to the death state and remained there for the rest of the simulation.

Figure 7: The cost-effectiveness model structure by health states



Model characteristics

Table 5: Features of the economic analysis

	Current appraisal	
Factor	**Chosen values**	**Justification**
Model structure	Individual patient-level micro-simulation	A patient-level simulation, rather than a cohort approach, was considered appropriate to capture changes in fracture risk, mortality and disease progression related to (re-) occurrence of fractures during the simulation. An individual state transition simulation model was considered to be more appropriate than a discrete event simulation (DES) approach. The main disadvantage of DES over an individual state transition in analysing the decision problem for this appraisal is the lack of accurate evidence for time-to-event data that would benefit a DES model, especially when event risks fluctuate over time due to imminent risks following fracture. This model structure is aligned with the ESCEO/IOF guidelines for osteoporosis modelling, and was previously validated under the NICE PRIMA process.[61-63]

Time horizon	Lifetime	Osteoporosis is a disease that affects patients for the remainder of their life. The model followed a patient from entering the model until death or age 100 years, whichever came first, in line with ESCEO/IOF guidelines.[63]
Cycle length	Six months	Sufficiently short to capture imminent fracture risk after fracture, and allowed for more than one transition with treatment effect for romosozumab. A six-month cycle length is aligned with ESCEO/IOF guidelines.[63]
Source of event probabilities	Clinical risk factors from ARCH[53] incorporated into a FRAX-based algorithm which additionally incorporated imminent risk from the Swedish registry[24, 60, 68]	FRAX currently underestimates the risk of imminent fracture in patients with a fragility fracture as it does not take into account predictors of imminent fracture (recency and site of fracture). Therefore, the FRAX-based algorithm used in the model includes these additional risk factors. The importance and impact on this on cost-effectiveness has been described in the literature.[64] Modelling increased risk after fracture events is aligned with ESCEO/IOF guidelines.[63]
Source of utilities	Fracture utility multipliers from ICUROS study* GIAE decrement from Davis *et al.* (2015)[69]	The ICUROS study was specifically designed to assess the QoL impact of fractures on osteoporosis over time with the objective of allowing the appropriate use of its findings in cost-effectiveness models. ICUROS captures the QoL impact of fracture as soon as possible after a fracture occurs, regardless of treatment. Conversely, QoL was assessed irrespective of fracture occurrence at predetermined discrete timepoints in the ARCH trial, and always in relation to one of the treatments investigated. It is therefore not appropriate to use the QoL data collected in ARCH because it does not provide robust health-related utility values that are sensitive to the decrease in QoL associated with fracture occurrence, and does not provide treatment-unspecific utility values required for valid economic evaluation. The independent academic Assessment Group used ICUROS in NICE TA464 and intended to do so again in the suspended NICE MTA ID901.[25, 70] The use of ICUROS utilities is also recommended by ESCEO/IOF. [63]
Source of costs	Romosozumab: UCB Comparators: BNF January 2021 drug tariff prices. Administration costs: derived from the relevant SmPC for each drug; GIAE-associated costs: Davis et al. (2015),[69] PSSRU[71], NHS Tariff Workbook 2020/21[72] Fracture costs: inflated from UK study by Gutiérrez et al. (2011 and 2012) using UK GP database[73, 74]	In accordance with the NICE reference case The acute costs of hip, vertebral and NHNV fractures were also used in the TA464[35] and were taken directly from the UK study by Gutiérrez et al. using UK GP database[73, 74]

Resource use	Acute costs: Study of postmenopausal women in the UK,[73, 74] Long-term: UK study of probability of discharge to institutional care by age[75]	In accordance with NICE reference case
Health effects measures	QALYs	In accordance with NICE reference case
Discount rate for costs and QALYs	3.5% per year	In accordance with NICE reference case
Perspective on costs	NHS and PSS	In accordance with NICE reference case

Abbreviations: BNF: British National Formulary; DES: discrete event simulation; eMIT: electronic market information tool; ESCEO: European Society for Clinical and Economic Aspects of Osteoporosis; FRAX: Fracture Risk Assessment Tool; GIAE: gastrointestinal adverse event; GP: general practitioner; ICUROS: International Costs and Utilities Related to Osteoporotic Fractures Study; IOF: International Osteoporosis Foundation; MTA: multiple technology appraisal; NA: not applicable; NHS: National Health Services; NHNV: non-hip, non-vertebral; NICE: National Institute for Health and Care Excellence; PRIMA: Preliminary Independent Model Advice; PSS: Personal Social Services; PSSRU: Personal Social Services Research Unit; QALY: quality-adjusted life year; QoL: quality of life; SmPC: Summary of Product Characteristics; TA: technology appraisal; UK: United Kingdom.

A.11 Incorporating clinical evidence into the model

Assessment of fracture risk

The risk of sustaining a fracture in the model was based on three elements: the risk for an individual in the general population incurring a fracture, the increased fracture risk associated with osteoporosis (the relative risk) and a risk reduction, if any, attributed to a treatment. Therefore, the risk of experiencing a fracture in the model is calculated as:

$$Age\ specific\ general\ population\ risk * Relative\ risk\ of\ fracture * Risk\ reduction\ from\ treatment$$

Increased fracture risk due to osteoporosis

The model employed in this submission used the FRAX algorithm to generate an estimated fracture risk. FRAX was used in the clinical trial setting of the ARCH study and NICE has concluded that cost-effectiveness results are broadly similar using FRAX and QFracture.[76]

FRAX is not currently capable of calculating the imminent risk as the current FRAX tool does not consider recency or site of prior fracture.[64] These factors are major drivers of another fracture, as described in Section A.1 . Therefore, the 10-year risk from FRAX will be an underestimation of the short-term fracture risk in patients who have experienced a recent fragility fracture and are at imminent risk of another fracture.[77] The importance and impact of this on cost-effectiveness have been described in the literature.[64]

In the model, whenever a patient sustained a fracture, their individual fracture risk was updated. Data to determine the increase in risk were taken from a retrospective real-world evidence study in Swedish women.[24, 60, 68] This study was used due to the lack of available data in the UK. Although estimates of absolute fracture values vary between countries, relative estimates can be assumed to be transferable across geographic settings. The data from Sweden are robust and

extensive, and therefore this database is the most reliable data source in Europe for osteoporosis.[60] Estimates from this study can be used to obtain estimates of absolute values for the UK, where no such data are available.

In this study, the imminent risk of another MOF was observed following the fracture in women with one, two or three fragility fractures.[60, 68] This risk decreased over time from index fracture, and little excess risk was observed after five years (adjusted for covariates). Therefore, it is possible to consider a "fracture cascade", whereby a fracture increases the short-term (imminent) risk of another fracture, which reduces over time (Figure 8). Imminent risk of another fracture provides supporting evidence that treatment intervention should be targeted as soon as possible after a fragility fracture.[11, 78] Romosozumab reaches the optimal clinical performance in a relatively short duration (i.e., 12 months), providing a rapid and potent effect and demonstrating the potential to interrupt such a "fracture cascade" early in the process. The approach in this model to include the fracture cascade was validated by a clinical expert during the NICE PRIMA review process, who described it as representing *"the classical progression of this disease from a state of increased fracture risk, to fracture, followed by an increased risk of another fracture which is highest in the months immediately after the index fracture"*.[62]

Fracture risk was estimated as a function of the UK general population risk, the RR estimated by FRAX for a given patient profile, and the maximum of the time-dependent RR of fracture:

$$MAX(RR_{recent\ fx\ versus\ no\ fx} \mid FRAX\ RR_{fx\ vs\ norm\ pop.}) * FRAX\ RR_{\ patient\ profile\ excl.fx\ CRF} * General\ population\ risk * Risk\ reduction\ from\ treatment$$

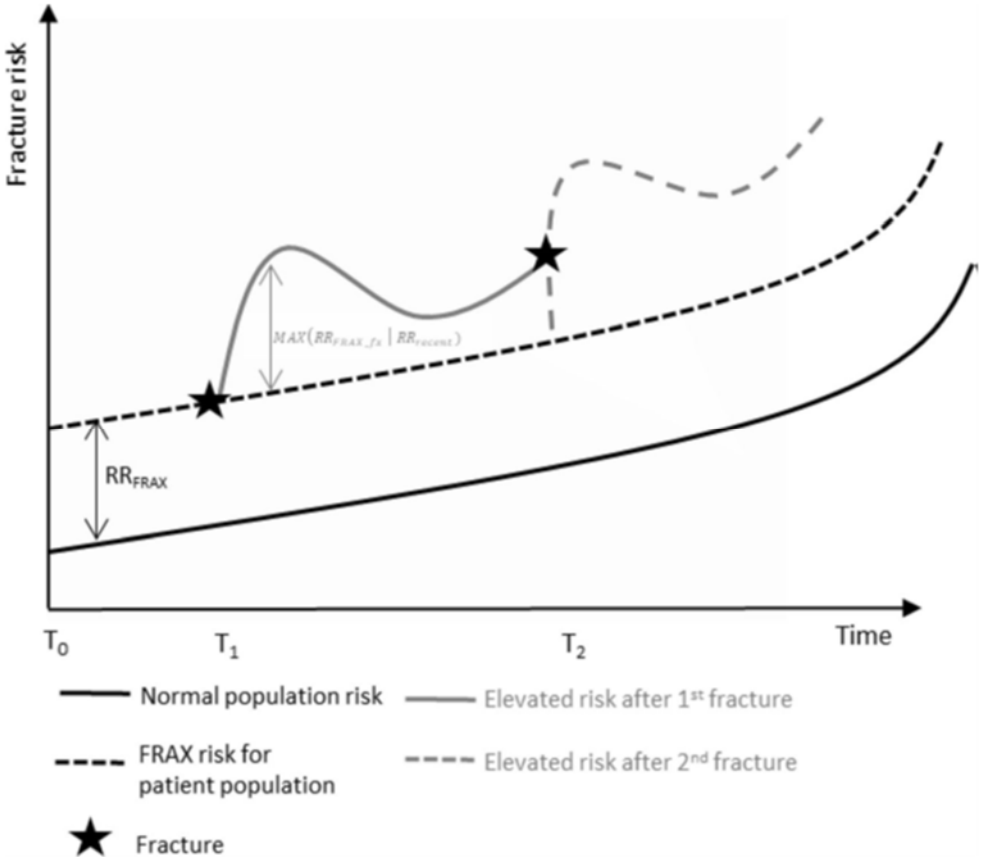
Abbreviations: CRF: clinical risk fracture; excl.: excluding; fx: fracture; RR: risk ratio; norm pop.: normal population.

Figure 8 provides an example of how the fracture risk trajectory was estimated at different time points in a patient without a fracture at baseline.

- T_0: At this point, the patient has no fracture history. The simulated fracture risk corresponds to the normal population's risk adjusted for the patient profile's clinical risk factors according to FRAX
- 1st Fracture: The patient suffers their first fracture. The simulated risk corresponds to the normal populations, risk-adjusted for the patient profile's clinical risk factors according to FRAX, and the maximum of time-dependent recent (1st) fracture RR and the RR of having fracture history according to FRAX
- 2nd Fracture: The patient suffers their second fracture. The simulated risk corresponds to the normal populations, risk-adjusted for the patient profile's clinical risk factors according to FRAX, and the maximum of time-dependent recent (2nd) fracture RR and the RR of having fracture history according to FRAX
- 3rd Fracture: The patient suffers their third fracture. The simulated risk corresponds to the normal populations, risk-adjusted for the patient profile's clinical risk factors according to FRAX and the maximum of time-dependent recent (2nd) fracture RR and the RR of having fracture history according to FRAX. Second fracture recent fracture RR was used because few patients experienced a third fracture in the source data, therefore, the RRs were associated with high uncertainty.

Risk related to recent fracture is not multiplicative.

Figure 8: How risk trajectory was estimated with imminent fracture risk in the model



Please note that this figure is for illustrative purposes only. **Source:** Söreskog et al. 2020[60]

Baseline Fracture Incidence

The model inputs for baseline incidences of hip, vertebral and NHNV fractures are summarised in Table 6. The incidences of hip fractures were based on a prospective study by Singer et al.[79] Although this article is from 1998, it has the currently most comprehensive data on hip fracture incidence in UK. A retrospective study using the Clinical Practice Research data link (CPRD) in the UK showed that fracture incidences have remained stable over the years 1990–2012 and similar to Singer et al.'s estimates,[80] which provides support on the use of data from the article by Singer et al.

Comprehensive data on the risk of clinical vertebral fractures in the UK are scarce. Although there are differences in incidences, the proportionality between fracture types is similar throughout the western world. Therefore, the UK clinical vertebral fracture incidence was calculated by assuming that the ratio of clinical vertebral fracture to hip fracture in a Swedish based study is similar to that of UK.[68] This method was also used in a report on osteoporosis in the European Union endorsed by IOF and the European Federation of Pharmaceutical Industry Associations (EFPIA).[81] NHNV fractures includes forearm (distal forearm, distal radius and wrist) and "other" fractures (femur, pelvis, humerus, rib, clavicle, scapula, sternum). Incidences of forearm fractures were taken from Singer et al.[79] Singer et al. have also published estimates of other fractures but did not report all fracture types (e.g. rib fractures). Therefore, the same imputation via hip fracture incidence and Swedish risk of "other fractures" was made for the combined incidence of "other fractures" in the UK.[79, 82]

Table 6: Incidence per 100,000 people in the UK by age

Age	Hip[79]	Vertebral[68]	NHNV[79, 82]
50–54	33	84	633
55–59	51	142	813
60–64	81	143	979
65–69	132	192	1,425
70–74	282	397	1,928
75–79	619	602	2,891
80–84	1,236	777	3,876
85+	2,255	1,061	5,958

Abbreviations: NHNV: non-hip, non-vertebral; UK: United Kingdom.
Source: Singer *et al.* (1998)[79], Kanis *et al.* (2000)[68], and Kanis *et al.* (2002)[82]

Risk reduction from treatment

The onset of treatment effect may vary by treatment. One of the benefits of romosozumab is that it is perceived to have a rapid onset of effect compared to other treatments, which may interrupt the "fracture cascade" in patients who are at high risk of fracture. The model applied efficacy estimates for romosozumab/alendronate and the comparators to the fracture risks of the patient population. In the base-case versus alendronate alone, the efficacy estimates were determined from the fracture endpoints from the ARCH study.[53]

ARCH is the only study of romosozumab in women with a prior fracture which includes fracture outcomes. Therefore, ARCH is the most relevant source of clinical evidence for modelling patients at imminent risk of fracture. Time-to-event analysis of fracture incidences are available from the Clinical Study Report (CSR) for clinical fracture, non-vertebral fracture, hip fracture, and MOF. Cumulative point estimates are published for 12 and 24 months for new vertebral, clinical, non-vertebral and hip fracture types.[53]

Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH. Patient level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-dependent hazard rates were calculated for the mid-point of the model cycle. In the model, efficacy of non-vertebral fractures was applied to NHNV fractures due to lack of data on all fractures excluding both hip and vertebral. For vertebral fractures, efficacy of new vertebral fractures was calculated from the published data at 12 and 24 months.[53] The resulting non-cumulative hazard ratios (HRs) of romosozumab/alendronate vs alendronate are described in Table 7.

Table 7: ARCH non-cumulative efficacy data based on parametric distributions. Hazard ratio of romosozumab/alendronate vs. alendronate by time point. ITT population.

Time since treatment start (months)	HR (hip fracture)	HR (new vertebral fracture, used for vertebral fracture in the model)	HR (non-vertebral fracture, used for NHNV fracture in the model)
0–6	[redacted]	[redacted]	[redacted]
7–12	[redacted]	[redacted]	[redacted]
13–18	[redacted]	[redacted]	[redacted]

19–24	[redacted]	[redacted]	[redacted]
25–30	[redacted]	[redacted]	[redacted]
31–36	[redacted]	[redacted]	[redacted]

Abbreviations: HR: hazard ratio; ITT: intention-to-treat; NHNV: non-hip, non-vertebral.

Transformation of ARCH trial data into romosozumab vs placebo, using alendronate vs. placebo from the NMA

The ARCH trial compares romosozumab/alendronate and alendronate; thus, no efficacy data vs placebo is available in the trial. In the model, fracture risk reductions from treatment are applied to the general population risk. Therefore, it is necessary to transform the ARCH efficacy of romosozumab vs alendronate to romosozumab vs no treatment. To calculate RRs for romosozumab/alendronate vs. no treatment, the HRs of romosozumab/alendronate vs. alendronate alone in Table 7 were applied to RRs of alendronate vs. placebo based on the NMA. As HRs (Table 7) and RRs (from the NMA, Section A.8) give practically the same information, it was deemed reasonable, given the lack of RR data from ARCH, to use these interchangeably. The approach of using the RRs of alendronate vs. placebo based on the NMA is reasonable given that the efficacy data of alendronate vs placebo from UCB's NMA do not differ significantly from other NMAs, for example NICE's most recent NMA.[70, 83]

The NMA provides efficacies for up to 36 months after treatment start. For all treatments with longer treatment durations, efficacy is extrapolated beyond 36 months until the end of the treatment duration. Table 8 presents the efficacy input of romosozumab/alendronate vs. placebo where efficacy has been calculated based on the NMA using the ITT population (used in the base case scenario).

Table 8: Fracture risk ratio (95% CI), by fracture type and time point of romosozumab/alendronate vs. placebo based on the ARCH trial and NMA (ITT populations)

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	42–48	[redacted]	[redacted]	[redacted]

Abbreviations: CI: confidence interval; ITT: intention-to-treat; NMA: network meta-analysis; NHNV: non-hip, non-vertebral.

A.12 Key model assumptions and inputs

The base-case patient characteristics inputs for the model are detailed in Table 9. The mean age, femoral neck T-score and BMI risk factors were chosen to align to the population from the ARCH trial.[53]

Table 9: Patient characteristics in the economic model base-case

Model parameter	Value	Source and appropriateness for modelling patient population in decision problem
Sex	Female	Licensed indication
Fracture history	Recent fracture (MOF within 24 months)	ARCH,[53] Swedish registry.[60] Specifying MOF aligns with the expected target population for romosozumab in clinical practice, to maximise the benefits of treatment
Mean age, years	74	ARCH[53]; comparable to the average age of postmenopausal women with osteoporosis in the UK[25, 32]
Mean femoral neck T-score (SD)	−2.90	ARCH[53]
Mean BMI	25.41	ARCH[53]
Mean 10-year MOF probability	30%	Target patient population

Abbreviations: BMI: body mass index; MOF: major osteoporotic fracture; SD: standard deviation; UK: United Kingdom.

The key model assumptions and inputs are given in Table 10.

Table 10: Key model assumptions and inputs

Component	Assumption	Justification
Fracture risk estimation approach (B.3.2.2, Page 67–69)	FRAX-based algorithm to include recency of fracture in the estimation of risk	The FRAX tool was selected as it is included in the National Osteoporosis Guideline Group's guideline[32] and its algorithm can be more easily adapted to consider the concept of imminent risk of fracture. NICE has concluded that cost-effectiveness results are broadly similar using FRAX and QFracture.[76] FRAX does not consider recency or site of prior fracture. These are major drivers of future fractures, and therefore the 10-year risk from FRAX will be an underestimation of the short-term fracture risk. As such, this submission incorporates recency and site of prior fracture alongside the FRAX algorithm. In the model, whenever a patient sustained a fracture, their individual fracture risk was updated. This approach was aligned with ESCEO/IOF guidelines.[63]
Discount rates for costs and effects (B.3.2.3, Page 69–73)	3.5%	In accordance with the NICE reference case
Modelling horizon (B.3.2.3, Page 69–73)	Lifetime, to a maximum age of 100 years	Osteoporosis is a disease that affects patients for the remainder of their life. The model followed a patient from entering the model until death or age 100 years, whichever came first, in line with ESCEO/IOF guidelines.[63]

Component	Assumption	Justification
GIAE modelling (B.3.4.4, Page 89)	Included as an average utility decrement at treatment start for 3% of patients treated with oral bisphosphonates	In line with the assumptions included Davis et al. (2015)[69] as part of NICE TA464[25]
Persistence (B.3.3.4, Page 82–84)	Included	Excluding persistence would overestimate treatment length and thereby efficacy (detailed in Document B, Section B.3.3.4)
Efficacy offset assumption (B.3.3.5, Page 84–85)	Dynamic offset equal to time on treatment	The time a patient remains on treatment is directly related to the expected duration of efficacy. Studies have suggested that alendronate, zoledronate and teriparatide have offset times similar to the treatment length, and there is no robust evidence to support differential offsets for other treatments, providing evidence for the dynamic model approach.[84-88] This was validated by leading UK clinical experts. A separate one-year fixed offset time was applied to denosumab, as the clinical effect is limited to within six months after stopping treatment.
Efficacy, romosozumab (B.3.3.3, Page 76–81)	ARCH trial combined with the NMA (ITT population), non-cumulative efficacy	Described in Section A.11
Efficacy, sequential romosozumab/ alendronate (B.3.3.3, Page 76–81)	ARCH trial combined with the NMA (ITT population), non-cumulative efficacy	Described in Section A.11
Efficacy, alendronate (B.3.3.3, Page 76–81)	ARCH trial combined with the NMA (ITT population) Cumulative efficacy 0–12 months for the first 12 months of treatment, 0–36 efficacy for the following periods	Described in Section A.11
Mortality (B.3.3.6, Page 85–87)	Mortality rates were comprised of three rates: age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor.	Fragility fractures are associated with significantly increased mortality.[20-23] It has been documented that patients with osteoporosis have a higher degree of frailty compared to the general population and that excess mortality after fragility fracture is not entirely attributable to the fracture event. A common assumption is that 30% of excess mortality is directly caused by fragility fracture.[89, 90] Therefore, it was assumed that 30% of excess mortality after hip, clinical vertebral or NHNV fracture was associated with the fracture event.[89, 90]

Abbreviations: BMI: body mass index; ESCEO: European Society for Clinical and Economic Aspects of Osteoporosis; FRAX: Fracture Risk Assessment Tool; GIAE: gastrointestinal adverse event; IOF: International Osteoporosis Foundation; ITT: intention-to-treat; MOF: major osteoporotic fracture; NHNV: non-hip, non-vertebral; NICE: National Institute for Health and Care Excellence; NMA: network meta-analysis; SD: standard deviation; TA: technology appraisal; UK: United Kingdom.

A.13 Base-case ICER (deterministic)

The base case cost-effectiveness results for romosozumab versus alendronate and romosozumab versus no treatment are presented in Table 11 using the patient access scheme (PAS) price for romosozumab. The base case results show that, at PAS price, romosozumab/alendronate is a cost-effective treatment sequence for postmenopausal women with severe osteoporosis who have experienced a MOF (within 24 months) and are at imminent risk of another fracture, when compared to alendronate or no treatment, with incremental cost-effectiveness ratios (ICERs) of £16,660 or £3,747, respectively, per QALY gained.

Table 11: Base-case results with PAS (deterministic) – B.3.7 (Page 98)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental. costs (£)	Incremental LYG	Incremental QALYs	Pairwise ICER (£/QALY)
ROMO/ALN	[redacted]	10.045	[redacted]				
ALN	[redacted]	10.014	[redacted]	[redacted]	0.031	[redacted]	£16,660
No treatment	[redacted]	9.993	[redacted]	[redacted]	0.051	[redacted]	£3,747

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab/alendronate.

A.14 Probabilistic sensitivity analysis

Probabilistic sensitivity analyses (PSA) were conducted by simultaneously sampling from estimated probability distributions of model parameters to obtain 1,000 sets of model input estimates. Distributional assumptions for the model parameters are described below.

- The unit costs of drugs were taken as given and not sampled in the model. All other cost parameters were sampled assuming a lognormal distribution and a standard error of 25% of the base-case value.
- The utility multipliers for hip, vertebral and NHNV fracture were sampled using a lognormal distribution with standard errors based on study data.
- Persistence to treatment and the proportion of patients going to long term care after a hip fracture was sampled assuming a beta distribution.
- Relative risks for treatment efficacy were sampled assuming normal distribution and standard errors based on the trials and/or NMA.

The incremental results from the probabilistic analyses for the comparison of romosozumab/alendronate versus alendronate are presented in Table 12. A scatter plot of incremental costs and QALYs for romosozumab/alendronate versus alendronate alone and no treatment is presented in Figure 9 and the cost-effectiveness acceptability curve for this analysis is shown in Figure 10. Given a willingness-to-pay threshold of £30,000 per QALY gained, the probability that romosozumab-to-alendronate is cost-effective at PAS price vs. alendronate is [redacted]% and vs. no treatment [redacted]%.

Table 12: Base-case results with PAS (probabilistic) – B.3.8 (Page 99)

Technologies	Mean total costs (£)	Mean total QALYs	Mean incremental. costs (£)	Mean incremental QALYs	Pairwise ICER (£/QALY)
ROMO/ALN	[redacted]	[redacted]			
ALN	[redacted]	[redacted]	[redacted]	[redacted]	£14,537
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	£3,952

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab/alendronate.

Figure 9: Scatterplot of probabilistic results with PAS – Document B, B.3.8.1, Figure 16, (Page 99)

[redacted]

Abbreviations: ALE: alendronate; PAS: patient access scheme; ROM: romosozumab.

Figure 10: Cost-effectiveness acceptability curve with PAS – Document B, B.3.8.1, Figure 17 (Page 100)

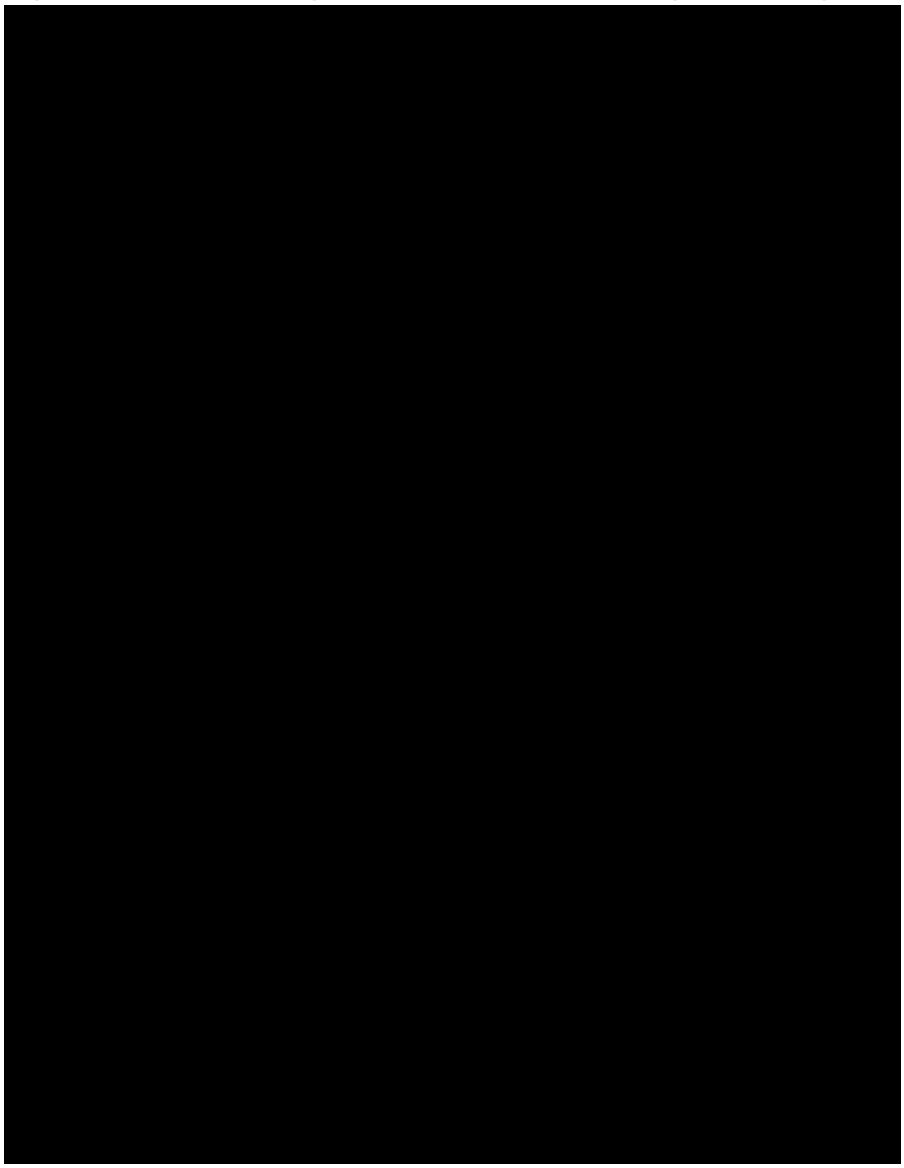
[redacted]

Abbreviations: ALE: alendronate; PAS: patient access scheme; QALY: quality-adjusted life year; ROM: romosozumab.

A.15 Key sensitivity and scenario analyses

Deterministic sensitivity analyses (DSA) were conducted by varying key parameters at lower and upper bounds of plausible values. Modelling assumptions were changed one-at-a-time to measure its impact on cost-effectiveness. A tornado plot summarising the DSA is presented in Figure 11

Figure 11: Tornado diagram with PAS – B.3.8.2 (Figure 18, Page 104)



Abbreviations: NHNV: non-hip, non-vertebral; RR: relative risk; PAS: patient access scheme.

Table 13 summarises the key scenario analyses and the pairwise ICER between romosozumab/alendronate and the chosen comparator in each scenario, compared to the pairwise ICER presented in the base case versus alendronate.

Table 13: Key scenario analyses with PAS

Scenario and cross reference	Scenario detail	Discussion	Pairwise ICER for ROMO/ALN vs comparator
Base case (ROMO/ALN vs ALN)			£16,660
Scenario 1 B.3.8.3, Page 104–108	12m romosozumab + 48m alendronate vs. 18m teriparatide biosimilar Movymia +alendronate 42m	Comparison with a sequence including the only bone-builder available in NHS practice (using biosimilar price for comparator)	ROMO/ALN Dominant
Scenario 2 B.3.8.3, Page 104–108	12m romosozumab + 48m alendronate vs. 18m teriparatide Forsteo +alendronate 42m	Comparison with a sequence including the only bone-builder available in NHS practice (using list price for comparator)	ROMO/ALN Dominant

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; PAS: patient access scheme; ROMO/ALN: romosozumab/alendronate.

A.16 Innovation

By inhibiting sclerostin, romosozumab allows activation of Wnt signalling that leads to a rapid increase in bone formation and BMD.[51] Romosozumab is the only dual-acting osteoanabolic biologic, with all other treatments being antiresorptives or a single-action anabolic. Antiresorptive therapies do not directly stimulate bone formation and therefore, romosozumab provides a clear advantage over bisphosphonates by rapidly increasing bone formation on naïve bone surface resulting in rapid improvements in bone density, mass, microstructure and strength leading to superior fracture risk reductions.[91, 92]

Romosozumab works rapidly, significantly reducing the incidence of new vertebral fractures by Month 12 versus alendronate. The relative risk of subsequent fracture is highest in the first two years following a fracture,[11] and therefore a treatment which can significantly reduce the risk of fracture over this time period will be beneficial in reducing the number of fractures experienced by patients with osteoporosis. It is likely that this rapid action on fracture reduction is the result of the rapid BMD improvements resulting from the mechanism of action: significant increases in BMD (vs alendronate) were observed as early as 6 months at lumbar spine, total hip and femoral neck in ARCH (see Section A.7.3).[56]

A.17 Budget impact

Table 14: Budget impact analysis results – Budget impact analysis template

	Company estimate				
	2022	**2023**	**2024**	**2025**	**2026**

Number of people in England estimated to receive treatment with romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Average annual treatment cost of romosozumab per person (treatment acquisition costs only) (PAS price)	[redacted]				
Estimated annual budget impact on the NHS in England (romosozumab at list price)	**2022**	**2023**	**2024**	**2025**	**2026**
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Estimated annual budget impact on the NHS in England (romosozumab at PAS price)	**2022**	**2023**	**2024**	**2025**	**2026**
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: PAS: patient access scheme; NHS: National Health Service.

A.18 Interpretation and conclusions of the evidence

Romosozumab is a unique osteoporosis biologic therapy with a dual-effect mechanism of action that acts to both stimulate bone formation and reduce bone resorption.[52-55] The results from the ARCH, FRAME and STRUCTURE trials provide evidence that romosozumab significantly reduces the incidence of new vertebral fractures and rapidly increases bone mass and strength. In the ARCH trial, treatment with romosozumab followed by alendronate resulted in a significantly lower risk of vertebral fracture at Month 24 versus women treated with alendronate alone in postmenopausal women with severe osteoporosis.

Similarly, positive results were observed in the FRAME trial, and were apparent in an NMA of the relative effectiveness of multiple osteoporosis treatments, which demonstrated that romosozumab was significantly more effective than or at least as good as most of the comparators included in the NMA. Romosozumab for 12 months followed by alendronate (romosozumab/alendronate) was the treatment [redacted] reported as the treatment with the [redacted] or [redacted] probability of being the most effective treatment.

Romosozumab therefore represents an important addition to the armamentarium for treatment of severe osteoporosis in postmenopausal women, with the potential to prevent fragility fractures and the associated pain, disability, detriment to HRQoL and mortality.[15-23]

The cost-effectiveness analysis demonstrated that using romosozumab first, before anti-resorptive therapy, is cost-effective versus anti-resorptive therapy alone when treating postmenopausal women who have experienced a MOF (within 24 months) and are at imminent risk of another fracture, equating to a modelled 10-year MOF FRAX probability of 30%.

In the base-case, romosozumab followed by alendronate is cost-effective at PAS price versus alendronate alone. Furthermore, romosozumab/alendronate is cost-effective versus no treatment. Scenario analyses demonstrated that romosozumab/alendronate at PAS price was dominant versus teriparatide (Forsteo® and biosimilar Movymia®) and cost-effective versus risedronate, zoledronate and raloxifene at a FRAX fracture probability of 30%.

References

1. Svedbom A, Hernlund E, Ivergard M, et al. Osteoporosis in the European Union: A Compendium of Country-Specific Reports. Arch Osteoporos 2013;8:137.
2. Tella SH, Gallagher JC. Prevention and treatment of postmenopausal osteoporosis. J Steroid Biochem Mol Biol 2014;142:155-70.
3. IOF. Key statistics for Europe. Available from https://www.osteoporosis.foundation/facts-statistics/key-statistic-for-europe [accessed 12 March 2021]. Volume 2018, 2017.
4. Bouxsein M. Biomechanics of osteoporotic fractures. Clinical Reviews in Bone and Mineral Metabolism 2006;4:143-153.
5. Recker R, Lappe J, Davies KM, et al. Bone remodeling increases substantially in the years after menopause and remains increased in older osteoporosis patients. J Bone Miner Res 2004;19:1628-33.
6. Seeman E. Structural basis of growth-related gain and age-related loss of bone strength. Rheumatology (Oxford) 2008;47 Suppl 4:iv2-8.
7. Brown JP, Josse RG, Scientific Advisory Council of the Osteoporosis Society of C. 2002 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada. CMAJ 2002;167:S1-34.
8. IOF. Facts and Statistics. Available from https://www.osteoporosis.foundation/facts-statistics [accessed 12 March 2021]. Volume 2018, 2017.
9. Kanis JA, Johansson H, Oden A, et al. Characteristics of Recurrent Fractures. Osteoporos Int 2018;29:1747–1757.
10. Adachi JD, Brown JP, Schemitsch E, et al. Fragility fracture identifies patients at imminent risk for subsequent fracture: real-world retrospective database study in Ontario, Canada. BMC Musculoskelet Disord 2021;22:224.
11. Johansson H, Siggeirsdóttir K, Harvey NC, et al. Imminent Risk of Fracture after Fracture. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2017;28:775–780.
12. van Geel TA, van Helden S, Geusens PP, et al. Clinical Subsequent Fractures Cluster in Time After First Fractures. Ann Rheum Dis 2009;68:99–102.
13. Pinedo-Villanueva R, Charokopou M, Toth E, et al. Imminent Fracture Risk Assessments in the UK FLS Setting: Implications and Challenges. Archives of Osteoporosis 2019;14:12.
14. Balasubramanian A, Zhang J, Chen L, et al. Risk of subsequent fracture after prior fracture among older women. Osteoporos Int 2019;30:79-92.
15. Dyer SM, Crotty M, Fairhall N, et al. A Critical Review of the Long-Term Disability Outcomes Following Hip Fracture. BMC Geriatrics 2016;16:158.
16. Kanis JA, Oden A, Johnell O, et al. The Burden of Osteoporotic Fractures: A Method for Setting Intervention Thresholds. Osteoporos Int 2001;12:417–427.
17. Cooper C, Jakob F, Chinn C, et al. Fracture incidence and changes in quality of life in women with an inadequate clinical outcome from osteoporosis therapy: the Observational Study of Severe Osteoporosis (OSSO). Osteoporos Int 2008;19:493-501.
18. Oglesby AK, Minshall ME, Shen W, et al. The impact of incident vertebral and non-vertebral fragility fractures on health-related quality of life in established postmenopausal osteoporosis: results from the teriparatide randomized, placebo-controlled trial in postmenopausal women. J Rheumatol 2003;30:1579-83.
19. Silverman SL, Minshall ME, Shen W, et al. The relationship of health-related quality of life to prevalent and incident vertebral fractures in postmenopausal women with osteoporosis: results from the Multiple Outcomes of Raloxifene Evaluation Study. Arthritis Rheum 2001;44:2611-9.
20. Cauley JA, Thompson DE, Ensrud KC, et al. Risk of Mortality Following Clinical Fractures. Osteoporos Int 2000;11:556–561.
21. Center JR, Nguyen TV, Schneider D, et al. Mortality after all major types of osteoporotic fracture in men and women: an observational study. Lancet 1999;353:878-82.
22. Katsoulis M, Benetou V, Karapetyan T, et al. Excess mortality after hip fracture in elderly persons from Europe and the USA: the CHANCES project. J Intern Med 2017;281:300-310.
23. Vestergaard P, Rejnmark L, Mosekilde L. Increased mortality in patients with a hip fracture-effect of pre-morbid conditions and post-fracture complications. Osteoporos Int 2007;18:1583-93.

24. Banefelt J, Åkesson K, Spångeus A, et al. Risk of imminent fracture following a previous fracture in a Swedish database study. Osteoporosis International 2019;30:601-609.
25. National Institute for Health and Care Excellence (NICE). Bisphosphonates for treating osteoporosis [TA464]. Available at https://www.nice.org.uk/guidance/ta464 [accessed 22 April 2021], 2019.
26. Fosamox (alendronate) SmPC. Available at: https://www.medicines.org.uk/emc/product/1281. Last accessed: 23 April 2021.
27. Aclasta (zoledronic acid) SmPC. Available at: https://www.medicines.org.uk/emc/product/210. Last accessed: 23 April 2021.
28. Actonel (risedronate sodium) SmPC. Available at: https://www.medicines.org.uk/emc/product/3835. Last accessed: 23 April 2021.
29. Bonviva (ibandronic acid) SmPc. Available at: https://www.medicines.org.uk/emc/product/117. Last accessed: 23 April 2021.
30. Bonviva (ibandronic acid solution). Available at: https://www.medicines.org.uk/emc/product/185. Last accessed: 23 April 2021.
31. Prolia (denosumab) SmPC. Available at: https://www.medicines.org.uk/emc/product/568. Last accessed: 23 April 2021.
32. National Osteoporosis Guidelines Group. NOGG 2017: Clinical Guideline for the Prevention and Treatment of Osteoporosis. Available at: https://www.sheffield.ac.uk/NOGG/NOGG%20Guideline%202017.pdf. Last accessed: 23 April 2021.
33. National Institute for Health and Care Excellence (NICE). Denosumab for the prevention of osteoporotic fractures in postmenopausal women [TA204]. Available at https://www.nice.org.uk/guidance/TA204 (accessed 23 April 2021). 2010.
34. Drake MT, Clarke BL, Khosla S. Bisphosphonates: mechanism of action and role in clinical practice. Mayo Clinic proceedings 2008;83:1032-1045.
35. National Institute for Health and Care Excellence (NICE). Raloxifene and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women [TA161]. Available at https://www.nice.org.uk/guidance/TA161 (accessed 23 April 2021). 2008.
36. Jonsson B, Strom O, Eisman JA, et al. Cost-Effectiveness of Denosumab for the Treatment of Postmenopausal Osteoporosis. Osteoporos Int 2011;22:967–982.
37. Johnell O, Kanis JA, Oden A, et al. Mortality after Osteoporotic Fractures. Osteoporos Int 2004;15:38–42.
38. Barrett JA, Baron JA, Beach ML. Mortality and Pulmonary Embolism after Fracture in the Elderly. Osteoporos Int 2003;14:889–894.
39. Schnell S, Friedman SM, Mendelson DA, et al. The 1-Year Mortality of Patients Treated in a Hip Fracture Program for Elders. Geriatric Orthopaedic Surgery & Rehabilitation 2010;1:6–14.
40. Morin S, Lix LM, Azimaee M, et al. Mortality Rates After Incident Non-Traumatic Fractures in Older Men and Women. Osteoporos Int 2011;22:2439–2448.
41. Bliuc D, Nguyen ND, Nguyen TV, et al. Compound risk of high mortality following osteoporotic fracture and refracture in elderly women and men. J Bone Miner Res 2013;28:2317-24.
42. Beaupre LA, Cinats JG, Jones CA, et al. Does functional recovery in elderly hip fracture patients differ between patients admitted from long-term care and the community? J Gerontol A Biol Sci Med Sci 2007;62:1127-33.
43. Forsteo (teriparatide) SmPC. Available at: https://www.medicines.org.uk/emc/product/2215. Last accessed: 23 April 2021.
44. Subbiah V, Madsen VS, Raymond AK, et al. Of mice and men: divergent risks of teriparatide-induced osteosarcoma. Osteoporos Int 2010;21:1041-5.
45. Vahle JL, Long GG, Sandusky G, et al. Bone Neoplasms in F344 Rats Given Teriparatide [rhPTH(1-34)] Are Dependent on Duration of Treatment and Dose. Toxicologic Pathology 2004;32:426-438.
46. Alswat KA. Gender Disparities in Osteoporosis. Journal of clinical medicine research 2017;9:382-387.
47. Bhimjiyani A, Neuburger J, Jones T, et al. The effect of social deprivation on hip fracture incidence in England has not changed over 14 years: an analysis of the English Hospital Episodes Statistics (2001-2015). Osteoporos Int 2018;29:115-124.
48. Patel R, Bhimjiyani A, Ben-Shlomo Y, et al. Social deprivation predicts adverse health outcomes after hospital admission with hip fracture in England. Osteoporos Int 2021.

49. Lang IA, Hubbard RE, Andrew MK, et al. Neighborhood Deprivation, Individual Socioeconomic Status, and Frailty in Older Adults. Journal of the American Geriatrics Society 2009;57:1776-1780.
50. Valentin G, Pedersen SE, Christensen R, et al. Socio-economic inequalities in fragility fracture outcomes: a systematic review and meta-analysis of prognostic observational studies. Osteoporosis International 2020;31:31-42.
51. Bandeira L, Lewiecki EM, Bilezikian JP. Romosozumab for the treatment of osteoporosis. Expert Opin Biol Ther 2017;17:255-263.
52. EVENITY (romosozumab) Summary of Product Characteristics. Available at https://www.medicines.org.uk/emc/product/10956/smpc [accessed 23 April 2021].
53. Saag KG, Petersen J, Brandi ML, et al. Romosozumab or Alendronate for Fracture Prevention in Women with Osteoporosis. N Engl J Med 2017;377:1417–1427.
54. Cosman F, Crittenden DB, Adachi JD, et al. Romosozumab Treatment in Postmenopausal Women with Osteoporosis. N Engl J Med 2016;375:1532–1543.
55. Langdahl BL, Libanati C, Crittenden DB, et al. Romosozumab (Sclerostin Monoclonal Antibody) Versus Teriparatide in Postmenopausal Women with Osteoporosis Transitioning from Oral Bisphosphonate Therapy: A Randomised, Open-Label, Phase 3 Trial. Lancet 2017;390:1585–1594.
56. UCB Data on File. ARCH Clinical Study Report - July 2017.
57. UCB Data on File. FRAME Clinical Study Report - June 2016.
58. UCB Data on File. STRUCTURE Clinical Study Report - November 2015.
59. Soreskog E, Lindberg I, Kanis JA, et al. Cost-effectiveness of romosozumab for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture in Sweden. Osteoporos Int 2021.
60. Söreskog E, Ström O, Spångéus A, et al. Risk of major osteoporotic fracture after first, second and third fracture in Swedish women aged 50 years and older. Bone 2020;134:115286.
61. National Institute for Health and Care Excellence. TA160: Raloxifene for the Primary Prevention of Osteoporotic Fragility Fractures in Postmenopausal Women. Available at: https://www.nice.org.uk/guidance/Ta160. Last accessed: 23 April 2021.
62. UCB Data on File. National Institute of Health and Care Excellence Scientific Advice - Preliminary Independent Model Advice and Clarification Questions. 2017.
63. Hiligsmann M, Reginster JY, Tosteson ANA, et al. Recommendations for the conduct of economic evaluations in osteoporosis: outcomes of an experts' consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the US branch of the International Osteoporosis Foundation. Osteoporos Int 2019;30:45-57.
64. Soreskog E, Borgstrom F, Lindberg I, et al. A novel economic framework to assess the cost-effectiveness of bone-forming agents in the prevention of fractures in patients with osteoporosis. Osteoporos Int 2021.
65. Tandvårds- och läkemedelsförmånsverket (The Dental and Pharmaceutical Benefits Agency). Underlag för beslut om subvention - Nyansökan [Decision for reimbursement], 2020.
66. Scottish Medicines Consortium. Medicines advice romosozumab (Evenity), 2020.
67. Kanis JA, Hans D, Cooper C, et al. Interpretation and Use of FRAX in Clinical Practice. Osteoporos Int 2011;22:2395–2411.
68. Kanis JA, Johnell O, Oden A, et al. Long-Term Risk of Osteoporotic Fracture in Malmo. Osteoporos Int 2000;11:669–674.
69. Davis, S. Bisphosphonates for Preventing Osteoporotic Fragility Fractures (Including a Partial Update of NICE Technology Appraisal Guidance 160 and 161) Assessment Report (2015). Available at: https://www.nice.org.uk/guidance/ta464/documents/osteoporosis-prevention-bisphosphonates-inc-part-rev-ta160-ta161-id782-assessment-report2. Last accessed: 23 April 2021.
70. National Institute for Health and Care Excellence (NICE). Assessment Group Report: Non-bisphosphonates for treating osteoporosis [ID901]. In: School of Health and Related Research (ScHARR) UoS, ed, 2019.
71. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care 2020. Available at https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/ [accessed 22 April 2021].

72. National Health Service (NHS) National Tariff Payment System. The National Tariff Workbook 2020/21. Available at https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/ [accessed 20 April 2021].
73. Gutiérrez L, Roskell N, Castellsague J, et al. Clinical Burden and Incremental Cost of Fractures in Postmenopausal Women in the United Kingdom. Bone 2012;51:324–331.
74. Gutiérrez L, Roskell N, Castellsague J, et al. Study of the Incremental Cost and Clinical Burden of Hip Fractures in Postmenopausal Women in the United Kingdom. Journal of Medical Economics 2011;14:99–107.
75. Nanjayan SK, John J, Swamy G, et al. Predictors of change in 'discharge destination' following treatment for fracture neck of femur. Injury 2014;45:1080-4.
76. National Institute for Health and Care Excellence. CG146. Osteoporosis: assessing the risk of fragility fracture. [accessed 23 April 2021] Available from: https://www.nice.org.uk/guidance/cg146/chapter/introduction.
77. Kanis J, Harvey N, McCloskey E, et al. Algorithm for the management of patients at low, high and very high risk of osteoporotic fractures. Osteoporosis International 2020;31:1-12.
78. Jonsson E, Ström O, Spångéus A, et al. Risk of Major Osteoporotic Fracture (Hip, Vertebral, Radius, Humerus [MOF]) After First, Second and Third Fragility Fracture In A Swedish General Population Cohort. Value in Health 2017;20:A528.
79. Singer BR, McLauchlan GJ, Robinson CM, et al. Epidemiology of Fractures in 15,000 Adults: The Influence of Age and Gender. J Bone Joint Surg Br 1998;80:243–248.
80. van der Velde RY, Wyers CE, Curtis EM, et al. Secular Trends in Fracture Incidence in the UK Between 1990 And 2012. Osteoporos Int 2016;27:3197–3206.
81. Hernlund E, Svedbom A, Ivergård M, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos 2013;8:136.
82. Kanis JA, Johnell O, Oden A, et al. Intervention Thresholds for Osteoporosis. Bone 2002;31:26–31.
83. Simpson EL, Martyn-St James M, Hamilton J, et al. Clinical effectiveness of denosumab, raloxifene, romosozumab, and teriparatide for the prevention of osteoporotic fragility fractures: A systematic review and network meta-analysis. Bone 2020;130:115081.
84. Black DM, Schwartz AV, Ensrud KE, et al. Effects of Continuing or Stopping Alendronate after 5 Years of Treatment: The Fracture Intervention Trial Long-term Extension (FLEX): A Randomized Trial. JAMA 2006;296:2927–2938.
85. Black DM, Reid IR, Boonen S, et al. The Effect of 3 Versus 6 Years of Zoledronic Acid Treatment of Osteoporosis: A Randomized Extension to the HORIZON-Pivotal Fracture Trial (PFT). J Bone Miner Res 2012;27:243–254.
86. Lindsay R, Scheele WH, Neer R, et al. Sustained Vertebral Fracture Risk Reduction After Withdrawal of Teriparatide in Postmenopausal Women with Osteoporosis. Arch Intern Med 2004;164:2024–2030.
87. Prince R, Sipos A, Hossain A, et al. Sustained Nonvertebral Fragility Fracture Risk Reduction After Discontinuation of Teriparatide Treatment. J Bone Miner Res 2005;20:1507–1513.
88. Strom O, Landfeldt E, Garellick G. Residual Effect After Oral Bisphosphonate Treatment and Healthy Adherer Effects--The Swedish Adherence Register Analysis (SARA). Osteoporos Int 2015;26:315–325.
89. Strom O, Borgstrom F, Sen SS, et al. Cost-effectiveness of alendronate in the treatment of postmenopausal women in 9 European countries--an economic evaluation based on the fracture intervention trial. Osteoporos Int 2007;18:1047-61.
90. Borgstrom F, Johnell O, Kanis JA, et al. Cost Effectiveness of Raloxifene in the Treatment of Osteoporosis in Sweden: An Economic Evaluation Based on the MORE Study. Pharmacoeconomics 2004;22:1153–1165.
91. Cosman F, Dempster DW. Anabolic Agents for Postmenopausal Osteoporosis: How Do You Choose? Curr Osteoporos Rep 2021;19:189-205.
92. Langdahl B, Ferrari S, Dempster DW. Bone modeling and remodeling: potential as therapeutic targets for the treatment of osteoporosis. Ther Adv Musculoskelet Dis 2016;8:225-235.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Romosozumab for treating severe osteoporosis [ID3936]

Clarification questions

June 2021

File name	Version	Contains confidential information	Date
ID3936 Romosozumab-clarification letter to PM for company	**1st draft**	**Yes**	**19/07/2021**

Notes for company

Highlighting in the template

Square brackets and [redacted] highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [redacted] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Literature Searches

A1. PRIORITY QUESTION: Please provide full details of the second update searches conducted for the clinical effectiveness systematic literature review (SLR) in September 2020 referred to in Appendix D.1.3.

[redacted]

A2. Please provide full search strategies for the clinical trial registries (ClinicalTrials.gov and WHO International Clinical Trials Registry Platform) searches in Appendices D.1.1. and D.1.2.

[redacted]

A3. Please provide full details of the searches of conference proceedings referred to in Appendices D.1.1. and D.1.2.

[redacted]

A4. Please provide full details of the searches of additional websites (health technology assessment organisations) referred to in Appendices D.1.1. and D.1.2.

[redacted]

Decision Problem

A5. PRIORITY QUESTION: Please clearly define a major osteoporotic fracture (MOF) and clarify whether this includes mild vertebral fractures as mentioned in the Clinical Study Report of the ARCH study (CSR, page 70).

[redacted]

A6. PRIORITY QUESTION: On page 43 of the company submission (CS) (Section B.2.7) it is stated that "The ARCH population is largely analogous to the proposed romosozumab target population, with the key difference being

that the ARCH trial did not mandate the prior fracture to be recent, whereas the romosozumab target population defines recency of fracture as a criterion."

A. **Please explain what proportion of patients in the ARCH trial 'experienced a recent major osteoporotic fracture (MOF) within 24 months', i.e. fulfilled the criteria for the population in the CS. Please provide exact numbers by treatment arm.**

B. **Please provide data from the ARCH trial for the subgroup of patients who 'experienced a recent MOF within 24 months' (i.e. the population in the CS)**

[redacted]

A7. Please clarify how easy it is to classify patients in UK practice as having severe osteoporosis (the population in the NICE scope) and as having severe osteoporosis and experienced a recent MOF within 24 months (the population in the CS).

[redacted]

Romosozumab Trials

A8. PRIORITY QUESTION: Section B2 (Summary of clinical evidence) of the CS explains that [redacted] were detected between romosozumab and comparators in the ARCH and FRAME trials (Health-related quality of life [HRQoL], page 26) however, then goes on to say that "While [redacted] were noted between treatment groups there were declines in HRQoL data following fractures on both treatments. By preventing fragility fractures, romosozumab is therefore expected to prevent the loss of HRQoL associated with fracture".

The above statement does not appear to follow logically from the preceding text. Please explain further or provide evidence to support the assertion that use of romosozumab is associated with maintenance of HRQoL.

[redacted]

A9. Section B2 (Summary of clinical evidence, Safety - page 27) of the CS states: "A numerical imbalance of incidence of myocardial infarction and stroke was noted in

the alendronate-controlled ARCH study". Please provide exact numbers by treatment arm.

[redacted]

A10. Section B2 (Summary of clinical evidence, Safety - page 27) of the CS states: "Across the trials, the most common adverse reactions were nasopharyngitis (13.6%) and arthralgia (12.4%). Hypersensitivity-related reactions occurred in 6.7% of patients treated with romosozumab. Hypocalcaemia was reported uncommonly (0.4% of patients treated with romosozumab)." However, these percentages are different from the percentages in Table 15 (page 56) of the CS. Please explain the differences, and please provide exact numbers and percentages per treatment group in all instances or signpost the reader to where this information can be found.

[redacted]

A11. The following information is stated as part of Section B.2 (Conclusion, page 27) of the CS: "Romosozumab is an important addition to the treatment armamentarium for postmenopausal women with severe osteoporosis, with the potential to prevent vertebral and non-vertebral fragility fractures and their associated pain, disability, detriment to HRQoL and mortality." Since pain, disability and mortality are not represented within the "Summary of clinical evidence" (Document B), the underpinning basis of the final statement is not clear. Furthermore, pain and disability are not outcomes specified for the submission. Please signpost the reader to the evidence underpinning each outcome mentioned in the above statement.

[redacted]

A12. Table 5 (Summary of methodologies for ARCH, Method of blinding) in Section B.2.3.1 of the CS states the following (and similar information appears in Section B.2.5, page 37): "Double-blind: patients and site staff remained blinded to the patient's original treatment assignment". Please explain how this was accomplished in light of romosozumab administration being by subcutaneous injection and alendronate being given orally.

[redacted]

A13. Section B.2.6 (Clinical effectiveness results of the relevant trials) of the CS includes this information: "The results from the ARCH trial presented in this section

describe those that were detailed in the ARCH Clinical Study Report (CSR) and were determined using the standard methodology of last observation carried forward (LOCF) imputation for missing data, as pre-specified in the statistical analysis plan. However, the data more recently presented in the peer-reviewed New England Journal of Medicine publication regarding fractures and bone mineral density (BMD) were determined using a multiple imputation for the missing data as requested by the journal, which does not reflect the original pre-specified analyses, and has thus not been included in this submission." Please clarify whether there were any differences in estimates of effect between the two methods of imputation, and describe how any differences between these analyses could affect the cost-effectiveness estimate.

[redacted]

A14. Regarding data extraction and quality assessment **(**Appendix D, Section D.2 of the CS), please clarify how disagreements about data extraction were resolved and please also clarify which version of the Cochrane Risk of Bias tool for RCTs was used.

[redacted]

A15. Please include 95% confidence intervals when reporting any and all effect estimates, relevant throughout, but particularly in Section B.2.6 of the CS. Also, please include 95% confidence intervals on bar charts presented in Figures 5, 6, 8 in Document B. Please also add the number of patients to Figure 8 in Document B, as in Figures 5 and 6 in Document B.

[redacted]

A16. Please further justify why the ARCH intention-to-treat (ITT) population is generalisable to the UK treatment population, including how many patients in ARCH were from the UK, and whether the demographics of the ARCH ITT population match that of the UK treatment population (particularly ethnic group and geographic region), and if not, how any differences are likely to affect the cost-effectiveness estimate.

[redacted]

A17. Please clarify whether in Table 6 of Document B "A 25-hydroxyvitamin D level of >20 ng/ml" should be "A 25-hydroxyvitamin D level of <20 ng/ml", given the median and 25th centile (in Table 7 of Document B) are both above 20 ng/ml for the ARCH trial population.

[redacted]

A18. Please justify why the per protocol analysis set was used for some outcomes rather than the ITT analysis set. Please also provide the ITT results for the incidence of new vertebral fractures and clarify for all analyses which analysis set is being used.

[redacted]

Indirect comparisons

A19. For all network meta-analysis (NMA) closed-loop analyses, please provide both the fixed-effect and random-effects estimates for the direct and indirect effects to placebo (for all treatments in the loop) so we can assess whether the inconsistency factors show whether the lack of statistically significant inconsistency is due to a lack of statistical power.

[redacted]

A20. Table 41 in Appendix D (page 143) is missing data for FRAME and Chao 2013, and Figure 16 in Appendix D is missing the Hadji 2012 study – please check all tables and figures in sections D.4.3, D.4.4 and D.4.5 to ensure all studies are included in the figures and in the tables and vice versa. Please also add percentages to the Events/N columns for all tables in these sections.

[redacted]

A21. PRIORITY QUESTION: The NMAs for BMD outcomes used the final time points of all included studies, unlike in the fracture outcomes where NMAs were specific to different time-points.

- **Please justify why the latest time-points were used, rather than splitting the NMAs into separate time points. If due to a lack of data, please justify this by showing the networks of evidence that would be for studies with outcomes at 12, 24 and 36 months.**

- **If feasible, please conduct separate NMAs using studies with outcomes at 12, 24 and 36 months for all BMD outcomes. If this is not feasible, please explain why and describe what effect combining different time-points may have on the results of the BMD NMAs.**
- **Please also add the time-point of analysis to all studies for all tables in Appendix D.4.5.**

[redacted]

A22. PRIORITY QUESTION: Please provide all analysis code for all analyses, including the WinBugs code and input data for the NMAs.

[redacted]

Section B: Clarification on cost-effectiveness data

Model structure and implementation

B1. PRIORITY QUESTION: Please define *all treatment sequences* included in the cost effectiveness analyses. This should include the base-case, the scenario analyses and the complete time horizon, *indicating also what effects are maintained and for how long*. This could be presented in the form of a table as below (please add rows/columns if needed):

[redacted]

Tx. arms	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11+
Intervention: ROMO + ALN	ROMO	ALN	ALN	ALN	ALN	?	?	?	?	?	?
Comp. 1: ALN	ALN	ALN	ALN	ALN	ALN	?	?	?	?	?	?
Comp. 2: Placebo											
Comp 3: TRP	TRP	TRP	?	?	?	?	?	?	?	?	?
…											

Abbreviations: ALN = alendronate, Comp. = comparator, ROMO = romosozumab, TRP = teriparatide, Tx. = treatment.

B2. On page 68 of the CS, it is mentioned that "All patients started the model in the "at risk" health state. At the end of each cycle patients either *moved* into the one of the fracture states, remained in their current health state without new fracture, or died." Please explain how transitions are determined in the model (e.g., by a random draw from different probability distributions).

[redacted]

B3. On page 66 of the CS, it is mentioned that “The algorithm used to generate the estimated fracture risk within the model is based on FRAX, but also includes an additional risk associated with recent fracture”. Please provide a numerical example illustrating how FRAX and the additional risk associated with recent fracture are used in the economic model. Please clarify whether this additional risk has been validated by experts. Finally, please conduct a scenario analysis based on FRAX only.

[redacted]

B4. PRIORITY QUESTION: On page 69 of the CS, there are several statements that require further clarification:

A. “FRAX is not currently capable of calculating the imminent risk as the current FRAX tool does not consider recency or site of prior fracture”. Please clarify whether other tools (similar to FRAX) are capable to calculate this imminent risk.

[redacted]

B. “Therefore, the 10-year risk from FRAX will be an underestimation of the short-term fracture risk in patients who have experienced a recent fragility fracture and are at imminent risk of another fracture”. Please provide an indication of the magnitude of the underestimation. Please explain how 10-year risks are converted into 6-month transition probabilities.

[redacted]

C. “In the model, whenever a patient sustained a fracture, their individual fracture risk was updated.” Please provide a numerical example illustrating how the individual fracture risk is updated.

[redacted]

D. “Although estimates of absolute fracture values vary between countries, relative estimates can be assumed to be transferable across geographic settings.” Please provide evidence to support this statement.

[redacted]

E. Please provide a numerical example illustrating how fracture risk is estimated in the economic model. Please describe all elements in the equation on page 69 of the CS.

[redacted]

Clinical parameters

B5. PRIORITY QUESTION: Please clarify the following questions regarding Table 17 of the CS:

A. Please clarify whether <u>all</u> patient characteristics are representative for UK patients. This is only indicated for age but not for the other characteristics.

[redacted]

B. Please justify the choice of 30% for the 10-year MOF probability.

[redacted]

C. Please explain how sensitive the model results are to changes in patient characteristics.

[redacted]

B6. PRIORITY QUESTION: Please clarify the following questions regarding <u>baseline fracture incidence</u>:

A. On page 78 of the CS it is mentioned that "A retrospective study using the Clinical Practice Research data link (CPRD) in the UK showed that fracture incidences have remained <u>stable</u> over the years 1990–2012 and <u>similar</u> to Singer et al.'s estimates". Please explain (numerically) to what extent fracture incidences have remained stable and similar to those in Singer et al. study.

[redacted]

B. On page 78 of the CS it is mentioned that "Comprehensive data on the risk of clinical vertebral fractures is limited for the UK, therefore, the UK clinical vertebral fracture incidence was calculated by assuming that the

ratio of clinical vertebral fracture to hip fracture in a Swedish-based study is similar to that of the UK":

- ***Please explain (numerically) to what extent data on the risk of clinical vertebral fractures is limited for the UK.***
- ***Please clarify why the study by Singer et al. has not been deemed appropriate to inform vertebral fractures but it was appropriate for hip and NHNV fractures.***
- ***Please indicate whether the assumption that the ratio of clinical vertebral fracture to hip fracture in a Swedish-based study is similar to that of the UK has been validated by clinical experts.***

[redacted]

C. Please explore scenario analyses where vertebral fractures are informed by Singer et al. and where the ratio of clinical vertebral fracture to hip fracture is changed in another (plausible) way.

[redacted]

B7. PRIORITY QUESTION: Please clarify the following questions regarding <u>risk reduction from treatment</u>:

A. On page 79 of the CS it is mentioned that "Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH". Please provide an example showing how the "continuous hazards approach" was applied.

[redacted]

B. On page 79 of the CS it is mentioned that "Patient level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-

dependent hazard rates were calculated for the mid-point of the model cycle". Please indicate where these analyses can be found.

[redacted]

C. Please explain how the hazard ratios shown in Table 19 were calculated. Please indicate what hazard ratios are used in the model after 36 months (until the end of the time horizon).

[redacted]

D. Please justify (both numerically and conceptually) why HRs (from Table 19) and RRs (from the NMA) "give practically the same information".

[redacted]

E. On page 79 of the CSit is mentioned that "The approach of using the alendronate vs. placebo data is reasonable given that the efficacy data of alendronate vs placebo from UCB's NMA do not differ significantly from other NMAs, for example NICE's most recent NMA (Table 20)". We consider this statement rather subjective seeing the values presented in Table 20. This is particularly the case for the values shown for teriparatide, which is the most effective treatment according to the AG NMA but not in the company's NMA. This raises concerns about the validity/credibility of the NMA results. Please provide separate results based on either NMA.

[redacted]

B8. Please provide new versions of Table 21 and 22 showing risk ratios for the complete modelled time horizon.

[redacted]

B9. PRIORITY QUESTION: Please clarify the following questions regarding <u>modelling of persistence</u>:

A. Please indicate the main causes for treatment discontinuation as observed in the ARCH trial and in the UK study by Li et al. 2010.

[redacted]

B. On page 84 of the CS it is mentioned that "Treatment discontinuation resulted in patients not receiving the same anti-fracture benefits as would be expected for a fully persistent patient (i.e., a patient still on treatment)". Please clarify whether the "effects" of treatment discontinuation have also been included in the costs and quality of life sides of the economic analyses.

[redacted]

C. Please justify the assumption that patients are at risk of dropping out during the first three years. Please clarify whether this assumption is applied to all treatments, regardless of the sequence. For example, for the intervention romosozumab (ROMO) + alendronate (ALN), patients are at risk of dropping out during the first three years in total (1 year of ROMO and 2 years of ALN) or during the first three years per treatment (1 year of ROMO and 3 years of ALN – so 4 years in total). In any case, this does not seem to match with the values shown in Table 25 where discontinuation is possible for some treatments up to year 5.

[redacted]

D. Please clarify what happens to patients after dropping out of one treatment: do they switch to the next in the sequence or do they all go to placebo? Please justify this assumption.

[redacted]

E. On page 84 of the CS it is mentioned that "In the base case cost-effectiveness analysis, persistence on alendronate alone (when not preceded by romosozumab) was derived from Li et al. (2012), a UK General Practice Research Database (GPRD) of persistence on osteoporosis medications among postmenopausal women in the UK". Please clarify why this was not based on ARCH data. Please provide a comparison between persistence estimates in Li et al. and the ARCH trial.

[Company: please enter your answer to this question here]

F. There are several sources of uncertainty regarding persistence on romosozumab in UK clinical practice and the company has acknowledged that this is still unknown. However, there are certain assumptions that require further justification:

i. **As a starting point, a Swedish study reporting persistence on teriparatide has been used. Please indicate whether it was not possible to use UK studies for this. In case it was not, please justify that the Swedish study is representative for the UK.**

[Company: please enter your answer to this question here]

ii. **The company stated that since romosozumab will be administered much less frequently compared to teriparatide, it is reasonable to assume that patients treated with romosozumab will exhibit higher persistence compared with teriparatide. While this might be the case, it might also be possible that patients could discontinue romosozumab for other reasons. Please justify this assumption.**

[Company: please enter your answer to this question here]

iii. **The magnitude of the improvement in persistence on romosozumab is unknown. The estimated persistence was estimated from clinical trial data. It might be expected that persistence is higher in clinical trials than in daily practice. That might be the reason why persistence on alendronate alone was derived from Li et al. (2012) and persistence on teriparatide was derived from the Swedish osteoporosis database. If that's the case, this approach (using trial data for romosozumab only) would be inconsistent and most likely biased in favour of romosozumab. Also, the assumptions made on page 85 of the CS "For the treatment sequence of romosozumab followed by alendronate used in this submission, <u>it was assumed that the persistence rates for alendronate were 85% of the persistence of denosumab</u>. This is based on the assumption that patients who have initially demonstrated high persistence on romosozumab would be expected to demonstrate high persistence**

on follow-on treatments, and therefore the persistence on alendronate after romosozumab would be notably higher than the persistence on alendronate alone reported by Li et al. (2012)"; are not justified enough. For those reasons, the estimates provided on Table 25 are uncertain, some of them inconsistent/unjustified and likely to favour romosozumab. Therefore, in any case, please conduct three additional scenarios, where:

- **persistence estimates for all treatments are based on trial data (even though this would most likely overestimate persistence for all treatments);**
- **persistence estimates for romosozumab are equal to persistence estimates for teriparatide (even though it might be expected that for romosozumab these would be higher – this could be seen as a conservative approach); and**
- **persistence is 100% (no treatment discontinuation).**

[Company: please enter your answer to this question here]

B10. PRIORITY QUESTION: Please clarify the following questions regarding dynamic residual effects:

A. Please provide numerical examples illustrating how dynamic residual effects are included in the model.

[Company: please enter your answer to this question here]

B. Please define also what is meant by "partially persistent patients" and include these patients in the numerical examples.

[redacted]

B11. PRIORITY QUESTION: Please clarify the following questions regarding modelling of mortality:

A. Please provide a numerical example showing how the three mortality rates mentioned in the CS (age-specific mortality of the general

population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor) are included in the model.

[redacted]

B. On page 87 of the CS it is mentioned that "All patients are at risk of dying corresponding to the risk of the UK general population from the start of the model". Please clarify why at the start of the model the risk of dying is not that of the patient population.

[Company: please enter your answer to this question here]

C. Please justify the choice of 30% relative risk of death associated to a fracture compared to no fracture (CS pages 87-88).

[Company: please enter your answer to this question here]

D. Please justify the assumption that "the standardised mortality ratios (SMRs) estimated using the Swedish data would be generalisable to the UK due to the similarity in access to health care between the two countries" (CS page 88). Please conduct scenario analyses where this SMR is varied within a plausible range of values.

[Company: please enter your answer to this question here]

E. On page 88 of the CS, it is mentioned that "As the variation in fracture distribution was not considered to be large across different age groups, the same relative risk was used for all ages". Please provide evidence to support this assumption. Also, please explain why "Using the same relative risk after NHNV fractures for all ages could thus possibly underestimate mortality in younger patients and overestimate mortality in older patients".

[Company: please enter your answer to this question here]

Adverse events

B12. PRIORITY QUESTION: P90 of the CS states "an imbalance in serious adjudicated cardiovascular (CV) adverse events (AEs) was observed in the

ARCH trial. As a result, romosozumab is contraindicated for patients with previous myocardial infarction or stroke. Given this contraindication, which was not an exclusion criterion in the ARCH trial, it was considered reasonable to exclude CV AEs from the economic analysis".

- **Please conduct an analysis showing the proportion of people who experienced a CV AE in the ARCH trial who had a history of myocardial infarction or stroke.**
- **Please include an option in the model to include CV AE according to the incidence in the ARCH trial and relevant disutilities and costs.**

[redacted]

B13. PRIORITY QUESTION: Please justify why only gastrointestinal adverse events (AEs) are included in the model and provide the option in the model to include all AEs at or above a 5% incidence threshold for either treatment arm for all Grade 3 or higher AEs.

[redacted]

Health-related quality of life

B14. PRIORITY QUESTION: Were the utility multipliers from the ICUROS study based on data from all countries in the dataset, a subset of countries or UK-specific? Please also justify your choice. If possible, please present UK-specific multipliers and include the option to use these in the model, if not already present.

[redacted]

B15. PRIORITY QUESTION: The ICUROS appears to include EQ-5D-3L data, EQ-VAS data and time trade-off (TTO) data. Please ensure that the multipliers included in the model are based only on EQ-5D-3L data.

[redacted]

B16. PRIORITY QUESTION: NICE TA464 (bisphosphonates for treating osteoporosis) also used utility multipliers from the ICUROS study, but the

multipliers differ from those presented in the CS. Please explain the difference in values.

[redacted]

B17. PRIORITY QUESTION: The CS states that the disutilities for multiple fractures are accounted for in a multiplicative approach. Please respond to the following points:

a) **Was it possible for individuals to receive more than 1 acute multiplier at the same time?**

b) **Did all patients enter the model with the full age-related general population utilities or were multipliers already applied to some patients?**

c) **Please consider how plausible it is that multiple prior fractures have the same relative impact on HRQoL in the long-term (e.g. 5+ years after occurrence), when a new fracture is experienced in the last year.**

d) **Please provide evidence that the included fracture types continue to affect HRQoL to the same extent 2 years, 5 years, 10 years and longer after occurrence. Please clarify that the model's assumptions regarding the length of time fractures are assumed to continue to affect utility and consider the plausibility of these assumptions. Please add the option in the model to reduce the duration of impact of chronic (2nd year+) multipliers, if a lifetime impact of such fractures has been assumed.**

e) **Please add the option in the model to assume a maximum disutility approach (whereby only 1 multiplier is applied, for the most impactful fracture at any point in time) or any other approach or amendments to the multiplicative approach that the company considers could appropriately capture the impact of multiple fractures, both acute (in the last year) and chronic (second or more years).**

[redacted]

B18. PRIORITY QUESTION: Page 43 of the CS states

"[redacted]

[redacted]**." Please provide the fracture utility decrements and multipliers which would be obtained from the ARCH HRQoL study and provide further justification as to why these are considered inappropriate**.

[redacted]

B19. Please explain how the QALY loss of 0.0075 for gastrointestinal adverse events was calculated.

[redacted]

Resource use and costs

B20. PRIORITY QUESTION: The analysis does not include administration costs for drugs that are administered via a subcutaneous injection, neither for romosozumab nor for the comparators in the scenario analyses. For romosozumab, the company justifies this by referring to their plans to set up a Patient Support Program (PSP) that includes homecare service, an adherence support program, and training of injection techniques.

- **Please provide more details regarding these plans and specify the costs of services and health care resource to the NHS and PSS that when the PSP is in place would be borne by the company instead.**
- **Please provide the option in the model to include drug administration costs (i.e. for subcutaneous injections) that are borne by the NHS and PSS when the PSP is not in place for romosozumab, as well as for the relevant comparators that are used in scenario analyses.**

[redacted]

B21. PRIORITY QUESTION: The costs during the first year following a fracture were sourced from Gutiérrez et al., 2011 for hip fractures and from Gutiérrez et al., 2012 for vertebral and non-hip-non-vertebral fractures. Gutiérrez et al., 2011 provide cost estimates both as total costs for patients who incurred a hip fracture as well as incremental costs of patients who incurred a hip fracture

relative to matched controls. Since the estimates reported by Gutiérrez et al. pertain to the cost year 2006/2007, the costs were inflated to 2019/2020.

- **Please confirm that the total (i.e. not the incremental) cost estimates from Gutiérrez et al. were used in the analysis for patients who had a fracture but not for those who did not have a fracture, and please justify the appropriateness of this approach.**

- **Please include the option in the model to use either the total costs, whilst applying these to both patients with and without a fracture correspondingly, and the incremental costs of patients who had a fracture relative to those who did not, with the latter only applied to patients who had a fracture.**

- **Please provide details regarding which cost estimates were used and which indices were used to inflate the costs of fractures, to clarify exactly how the cost estimates used in the analysis were arrived at.**

- **Please justify the appropriateness of including rehabilitation costs only for hip fractures and not for other types of fractures. Please provide the option in the model to either include rehabilitation costs for all types of fractures for which these are relevant or exclude rehabilitation costs for all types of fractures.**

- **Please comment on the suitability of the hip fracture cost shown in Table 33 of the CS (£13,203), which is considerably higher than the cost used by the Assessment Group in NICE ID901 (£8,568; shown in Table 8 of the Assessment Report).**

[redacted]

B22. PRIORITY QUESTION: Please provide details regarding how the annual drug and management costs that are listed in Table 31 of the CS were calculated.

[redacted]

B23. PRIORITY QUESTION: Please explain whether the treatment costs as applied in the model are in line with treatment adherence as observed in the treatment effectiveness results that are used to inform the model, and provide the option in the model to apply treatment costs in line with data on adherence (e.g. as provided in Table 25 in the CS) for all treatments considered in both base case and scenario analyses.

[redacted]

B24. Please provide the rationale and functionality of the 'Morbidity cost shares' inputs on the 'Cost input' sheet of the model that is commented as an optional input.

[redacted]

B25. Please justify the appropriateness of assuming the costs of chemotherapy intravenous infusion for the administration of zoledronate.

[redacted]

Cost effectiveness analyses

B26. PRIORITY QUESTION: Please provide a detailed explanation for the results of scenarios that demonstrate a large impact on the cost-effectiveness results when alternative values or assumptions are used, including start age and time horizon.

[redacted]

Model validation

B27. Please provide a comparison of the distribution of fractures in the source data vs. the distribution of fractures in the simulation. The idea is to validate the statement on page 70 of the company submission "few patients experienced a third fracture in the source data".

[redacted]

Section C: Textual clarification and additional points

C1. Please correct the errors (#N/A and #NUM!) in the model ‘PSA input’ sheet.

[redacted]

C2. The macros included in the model are inside a password-protected VBA project.

a) Please provide the password for the VBA project.

b) Please provide a detailed explanation of the functionality and implementation for each macro included in the model.

[redacted]

C3. PRIORITY QUESTION: Please include in the model ‘Main settings’ sheet the option to select all comparators included in the analyses.

[redacted]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Romosozumab for treating severe osteoporosis [ID3936]

Clarification questions

August 2021

File name	Version	Contains confidential information	Date
ID3936 Romosozumab-clarification letter response	1st draft	Yes	02/08/2021

Section A: Clarification on effectiveness data

Literature Searches

A1. PRIORITY QUESTION: Please provide full details of the second update searches conducted for the clinical effectiveness systematic literature review (SLR) in September 2020 referred to in Appendix D.1.3.

The full details of the second update searches are detailed in Appendix A1 of the appendices to the clarification questions.

A2. Please provide full search strategies for the clinical trial registries (ClinicalTrials.gov and WHO International Clinical Trials Registry Platform) searches in Appendices D.1.1. and D.1.2.

The full details of the search strategies for the clinical trial registries are available in Appendix A2 of the appendices to the clarification questions.

A3. Please provide full details of the searches of conference proceedings referred to in Appendices D.1.1. and D.1.2.

The full details of the searches of conference proceedings are available in Appendix A of the appendices to the clarification questions.

A4. Please provide full details of the searches of additional websites (health technology assessment organisations) referred to in Appendices D.1.1. and D.1.2.

For all additional websites of different HTA bodies (i.e. Canadian Agency for Drugs and Technologies in Health [CADTH], European Medicines Agency/Committee for Medicinal Products for Human Use [EMA/CHMP], NICE, National Institute for Health Research [NIHR], Food and Drug Administration [FDA]), the following intervention search terms were used to identify relevant submissions/assessments.

- romosozumab
- teriparatide
- alendronate
- risedronate
- ibandronate
- zoledronate/zoledronic acid
- denosumab
- raloxifene
- abaloparatide

Decision Problem

A5. PRIORITY QUESTION: Please clearly define a major osteoporotic fracture (MOF) and clarify whether this includes mild vertebral fractures as mentioned in the Clinical Study Report of the ARCH study (CSR, page 70).

The clinical definition of a MOF is a fracture of hip, clinical vertebral, forearm or humerus.[1] The use of the term clinical vertebral is to separate out non-symptomatic vertebral fractures (which can be identified with radiological methods) from those that are symptomatic (i.e. clinical). All clinical vertebral fractures are included regardless of their grade (severe, moderate or mild fracture). In some circumstances MOF also includes pelvic fractures as these carry similar risk and morbidity as hip fractures.

In the context of clinical trials, MOF include all vertebral fractures (i.e. those identified radiologically irrespective of the presence or absence of symptoms). A clinical trial imposes the use of regularly scheduled radiological methods designed to capture all vertebral fractures and leads to the identification of radiological vertebral fractures that may not be identified in routine clinical practice, as such an intense radiological schedule is not the regular practice in the clinical setting. Accordingly, in the ARCH trial the analysis of MOF included all vertebral fractures (mild, moderate or severe).

Vertebral fractures, regardless of severity, are known to significantly increase the risk of further fractures and therefore asymptomatic fractures are considered a major risk factor for future fragility fractures.[2] In the UK, both asymptomatic and clinical fractures are to be inputted into the FRAX tool, as they are considered to carry the same subsequent fracture risk .[3]

In the cost-effectiveness model mild, moderate and severe vertebral fractures are included in fracture types grouped as MOF.

A6. PRIORITY QUESTION: On page 43 of the company submission (CS) (Section B.2.7) it is stated that "The ARCH population is largely analogous to the proposed romosozumab target population, with the key difference being that the ARCH trial did not mandate the prior fracture to be recent, whereas the romosozumab target population defines recency of fracture as a criterion."

A. Please explain what proportion of patients in the ARCH trial 'experienced a recent major osteoporotic fracture (MOF) within 24 months', i.e. fulfilled the criteria for the population in the CS. Please provide exact numbers by treatment arm.

Within the licensed indication for romosozumab, the target patient population considered in this submission are patients who have experienced a recent major osteoporotic fracture (MOF) within the past 24 months; and thus, are at imminent risk of another fragility fracture.

In the ARCH trial, a total of [redacted] patients had suffered a fracture within 0–24 months before randomisation ([redacted] in romosozumab/alendronate group; [redacted] in alendronate alone group). Of these, [redacted] patients in the romosozumab/alendronate group and [redacted] patients in the alendronate alone group suffered a recent MOF and

would be eligible for treatment with romosozumab according to the target patient population considered in this submission.

B. Please provide data from the ARCH trial for the subgroup of patients who 'experienced a recent MOF within 24 months' (i.e. the population in the CS)

UCB conducted a *post hoc* analysis of the primary endpoint in the ARCH trial (incidence of new vertebral fractures through Month 24), investigating a subgroup of patients with a recent MOF (0–24 months before randomisation). *Post hoc* subgroup analysis results should be interpreted with caution, as there was no stratification of randomisation in the trial on whether patients enrolled with or without a recent MOF.

The *post hoc* analysis found the [redacted] for new vertebral fracture through Month 24, demonstrating that the treatment effect in the subgroup of patients in the primary analysis set with a MOF in the preceding 24 months was not different to the treatment effect in the subgroup whose preceding MOF occurred greater than 24 months before randomisation [redacted]. By Month 24, [redacted] of patients with a recent MOF treated with romosozumab/alendronate in the primary analysis set experienced a new vertebral fracture, compared to [redacted] of patients treated with alendronate/alendronate. In comparison, in the intention-to-treat (ITT) population, [redacted] of patients treated with romosozumab/alendronate experienced a new vertebral fracture at Month 24, compared with [redacted] treated with alendronate/alendronate. Please note, patient numbers are different in the *post hoc* analysis, as patients had to have both a baseline and a follow up spinal X-ray.

There is consistency in the clinical outcomes and baseline characteristics between the ITT population and the subgroup of patients with recent fracture.

In conclusion, the ITT population results are generalisable to the target population of romosozumab.

A7. Please clarify how easy it is to classify patients in UK practice as having severe osteoporosis (the population in the NICE scope) and as having severe osteoporosis and experienced a recent MOF within 24 months (the population in the CS).

Based on the World Health Organisation (WHO), the definition of severe osteoporosis is a patient who presents with a bone mineral density (BMD) value below a T-score of −2.5 standard deviations (SDs) and has one or more fragility fractures (i.e. low impact fractures sustained from standing height or less). Although it is relatively straightforward to detect a recent clinical fragility fracture as a patient presents in the clinic, the diagnosis of osteoporosis can be overlooked and therefore the UK has pioneered services and guidelines to avoid misdiagnosing these patients at their most critical time.

Patients with severe osteoporosis and imminent risk of fracture are picked up via the Fracture Liaison Service (FLS) who identifies patients aged 50 years and older with a new fragility fracture (clinical or asymptomatic) and follows up with them for further bone assessments (i.e. a dual energy x-ray absorptiometry (DXA) scan to obtain a T-score, initiates treatment).[4] This is done in-hospital, out-of-hospital and GP settings via multiple routes and aims to identify all patients from the following groups:[5]

- Managed as inpatients on acute orthopaedic/trauma wards
- Managed as inpatients on general medical/care of the older person wards not requiring surgical fixation (e.g., pelvic, upper limb, acute vertebral fracture presentations)

- Presenting acutely and not requiring hospital admission but managed as outpatients via orthopaedic / emergency medicine fracture clinics
- Presenting acutely but not requiring hospital admission or fracture clinic follow-up
- Vertebral fractures newly identified on radiology reports (incidental or anticipated)
- New fractures as a result of a fall during a hospital stay
- Patients who fracture whilst away from home and present later to local orthopaedic or primary care services

One of the key performance indicators (KPIs) for the FLS, as per the British Orthopaedic Association Standard for Trauma, is the provision of a multifactorial bone health assessment within 3 months or less of the incident of fracture, which is in alignment with imminent risk.[6] The national clinical audit run by the Falls and Fragility Fracture Audit Programme (FFAP) collects data from the UK FLS units on behalf of the NHS and feeds back to the clinics to improve their services. So far at its peak, the national average for FLS services in the UK was the identification of 44.7% fragility fractures compared to estimated case load and 66.5% underwent an FLS assessment in less than 90 days from fracture.[7] The provision of an FLS service has been proven to increase the initiation of treatment in those most at need by more than three times compared to usual care and is a cost-effective strategy for the prevention of further fractures.[8]

The UK osteoporosis community has pioneered the use of improvement tools and service models to identify patients at risk of secondary fragility fracture and to stratify these as high risk and high imminent risk, such as the development of the National Osteoporosis Guideline Group (NOGG) clinical guideline.[1, 9]

The NOGG guideline is currently being updated to include new therapies such as romosozumab and update intervention thresholds to low (no treatment initiation, lifestyle change), high (initiation of antiresorptive) and very high or imminent risk of fracture (initiation of anabolic therapy first).[10] The level of risk identified for very high or imminent risk of fracture was comparable to that of women enrolled in trials of anabolic agents such as romosozumab.[11]

In the near future, the identification of imminent fracture risk (IFR) in patients with severe osteoporosis will become even simpler as new technology and clinical guidance become available. For example, the NHS is prioritising the identification of undiagnosed vertebral fractures as they have recently funded £36 million in a range of state-of-the-art AI technology to transform the quality of care and the speed of diagnoses for conditions such as osteoporosis. One project aims to analyse existing CT scans to identify undiagnosed or asymptomatic vertebral fractures to ensure osteoporosis is managed and treated.[12] And a recommendation has recently been released to ensure any radiological image including the spine, regardless of the indication for the study, is appropriately processed to identify vertebral fractures. The NICE clinical guideline CG146 ‘Osteoporosis: assessing the risk of fragility fracture’ also is currently being updated to include the latest evidence and has an expected publication date of 21/02/2024.[13]

Romosozumab Trials

A8. PRIORITY QUESTION: Section B2 (Summary of clinical evidence) of the CS explains that [redacted] were detected between romosozumab and comparators in the ARCH and FRAME trials (Health-related quality of life [HRQoL], page 26) however, then goes on to say that “While [redacted] were noted between treatment groups there were declines in HRQoL data following

fractures on both treatments. By preventing fragility fractures, romosozumab is therefore expected to prevent the loss of HRQoL associated with fracture".

The above statement does not appear to follow logically from the preceding text. Please explain further or provide evidence to support the assertion that use of romosozumab is associated with maintenance of HRQoL.

HRQoL data were available from ARCH and FRAME; in both studies, [redacted]. An unusual design feature of both trials was the collection of HRQoL data monthly for three months following a fracture (while otherwise, HRQoL data was collected once every six months in ARCH, for example). When comparing HRQoL before and after a fracture it was evident that fractures caused a significant loss in patient HRQoL that was still present after 3 months. [redacted], fragility fractures were reported to have a considerable detrimental impact on HRQoL. For example, the least squares (LS) mean change from the pre-fracture baseline Osteoporosis Assessment Questionnaire Short Version (OPAQ-SV) score [redacted] in both treatment groups in the ARCH trial, indicating [redacted] at each time point (Month 1, 2 and 3) after fracture. Similar [redacted] were observed across the HRQoL outcomes assessed in the ARCH trial (please see Section 11.1 of the ARCH CSR for further details on patient reported outcomes [PRO] results). Thus, it is clear that each fracture had a measurable negative impact on QoL and fewer fractures in one treatment group imply less reduction in the QoL. However, at the trial population level this improvement is diluted by the null effect from all those subjects who did not experience a fracture.

Evidence from the literature also document that fractures have a detrimental impact on QoL. One HRQoL survey found that 80% of older women would rather be dead than experience the loss of independence and QoL that results from a hip fracture and subsequent admission to a nursing home, valuing nursing home admission at 0.05 on a scale of 0–1, where death is equal to 0.[14]

The results of the ARCH and FRAME trials demonstrate that romosozumab significantly reduced the incidence of fractures compared to alendronate and placebo, respectively. By reducing the incidence of fragility fractures compared to the treatments currently used in clinical practice, it is therefore logical to conclude that the use of romosozumab would result in preventing the associated loss of HRQoL typically seen following a fracture compared to the currently used treatments in clinical practice. The pivotal trials were not designed and powered to demonstrate this anticipated favourable HRQoL outcome directly.

Further, it is important to consider why QoL data collected in ARCH and FRAME trials showed no differences at each regular assessment timepoint. In ARCH, QoL data were collected at predetermined, discrete time points (once every six months initially) irrespective of fracture occurrence during the trial. The fractures that occurred during the study were spread across the duration of the study and so at any individual timepoint the reduction in QoL in patients with a fracture was diluted by the large number of patients who did not experience a fracture at the discretionary timepoints evaluated. As a result, the QoL assessments at a specific timepoint underestimates the impact of the therapy.

A9. Section B2 (Summary of clinical evidence, Safety - page 27) of the CS states: "A numerical imbalance of incidence of myocardial infarction and stroke was noted in

the alendronate-controlled ARCH study". Please provide exact numbers by treatment arm.

The number of patients that experienced myocardial infarction and stroke in each treatment arm is presented in Table 1. As can be observed, the absolute difference in MI and strokes between romosozumab and alendronate during the 12-month blind period is less than 0.5%.

Table 1: Patient incidence of treatment-emergent serious adverse events by preferred term (≥ 0.5% patient incidence in any treatment group) (safety analysis set at the time of the primary analysis)

System Organ Class Preferred Term	Double-blind Period		Primary Analysis Period	
	Alendronate 70 mg QW (N = 2014) n (%)	Romo 210 mg QM (N = 2040) n (%)	Alendronate 70 mg QW/ Alendronate 70 mg QW (N = 2014) n (%)	Romo 210 mg QM/ Alendronate 70 mg QW (N = 2040) n (%)
Cardiac disorders	[redacted]	[redacted]	[redacted]	[redacted]
Acute myocardial infarction	[redacted]	[redacted]	[redacted]	[redacted]
Cardiac failure	[redacted]	[redacted]	[redacted]	[redacted]
Atrial fibrillation	[redacted]	[redacted]	[redacted]	[redacted]
Cardiac failure congestive	[redacted]	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]
Cerebrovascular accident	[redacted]	[redacted]	[redacted]	[redacted]
Transient ischaemic attack	[redacted]	[redacted]	[redacted]	[redacted]
Syncope	[redacted]	[redacted]	[redacted]	[redacted]
Ischaemic stroke	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: mg: milligram; QM; once monthly; QW: once weekly.

A10. Section B2 (Summary of clinical evidence, Safety - page 27) of the CS states: "Across the trials, the most common adverse reactions were nasopharyngitis (13.6%) and arthralgia (12.4%). Hypersensitivity-related reactions occurred in 6.7% of patients treated with romosozumab. Hypocalcaemia was reported uncommonly (0.4% of patients treated with romosozumab)." However, these percentages are different from the percentages in Table 15 (page 56) of the CS. Please explain the

differences, and please provide exact numbers and percentages per treatment group in all instances or signpost the reader to where this information can be found.

The apparent discrepancy between these tables arises because they present two different analyses of the safety data:

The results found on page 27 of the CS represent the most frequent (≥5.0% in total romosozumab or placebo groups) adverse events by preferred term in the 12-month placebo-controlled osteoporosis safety analysis set. This can be found in Table 10, page 51 of the "Romosozumab Integrated Summary of Safety" which can be found in the reference pack to this submission.

The results in Table 15 (page 56) of the CS are exposure-adjusted, and represent exposure-adjusted incidence rates of the most frequent (≥5.0 per 100 subject-years in total romosozumab or integrated control groups) adverse events by preferred term in the osteoporosis safety analysis set. This can be found in Table 12, page 56 of the "Romosozumab Integrated Summary of Safety" which can be found in the reference pack to this submission.

A11. The following information is stated as part of Section B.2 (Conclusion, page 27) of the CS: "Romosozumab is an important addition to the treatment armamentarium for postmenopausal women with severe osteoporosis, with the potential to prevent vertebral and non-vertebral fragility fractures and their associated pain, disability, detriment to HRQoL and mortality." Since pain, disability and mortality are not represented within the "Summary of clinical evidence" (Document B), the underpinning basis of the final statement is not clear. Furthermore, pain and disability are not outcomes specified for the submission. Please signpost the reader to the evidence underpinning each outcome mentioned in the above statement.

As detailed in Section B.1.3.1, it is widely reported in the published literature that fragility fractures result in considerable disability and pain, as well as significant impairments in mobility, reduced independence and increased frailty.[15-19] Fewer than half of all individuals who experience a hip fracture will be able to walk unassisted, and most will never return to the same mobility level as prior to the fracture.[20]

Fragility fractures are also associated with significantly increased mortality.[21-24] Patients with a hip or vertebral fracture are at approximately four or three times higher risk of death, respectively, in the first year following the fracture, when compared to those without a fracture.[25] Non-hip/non-vertebral (NHNV) fractures are associated with up to 20% excess mortality compared to the general population in the first five years following a fracture.[26-29]

Consequently, while romosozumab did not appear to have an impact on pain, disability or mortality, romosozumab indirectly impacts each of these outcomes by reducing the incidence of fractures (which are the cause of the pain, disability and mortality) compared to the currently used treatments in UK clinical practice, as evidenced by the data presented in Document B, Section B.2. The studies were not powered to detect those differences. However, by reducing the incidence of fragility fractures, it is therefore reasonable to conclude that a population of patients treated with romosozumab will experience a reduced level of pain, disability and mortality, relative to patients treated with currently available treatments, because these patients will

experience fewer fragility fractures compared to patients treated with currently available treatments.

A12. Table 5 (Summary of methodologies for ARCH, Method of blinding) in Section B.2.3.1 of the CS states the following (and similar information appears in Section B.2.5, page 37): “Double-blind: patients and site staff remained blinded to the patient’s original treatment assignment”. Please explain how this was accomplished in light of romosozumab administration being by subcutaneous injection and alendronate being given orally.

The double-blind nature of the trial was preserved through the use of matched placebos. Romosozumab was presented in a single-use 1 mL prefilled syringe as a sterile, clear colourless and preservative-free liquid containing 70 mg of romosozumab per mL. Patients in the alendronate alone group received an injectable placebo in place of romosozumab, which was presented in identical containers and stored/packaged the same as romosozumab.

Blinded alendronate was commercially manufactured and labelled and distributed using Amgen clinical study drug distribution procedures. Patients in the romosozumab/alendronate group received a placebo of alendronate during the double-blind period, which was presented in identical containers and stored/packaged the same as alendronate. Additional information on methods of blinding can be found in Section 8.4.2 of the ARCH CSR.

During the open label alendronate period where all patients received alendronate the patients and the sites remained blinded to the original randomisation treatment arm.

A13. Section B.2.6 (Clinical effectiveness results of the relevant trials) of the CS includes this information: “The results from the ARCH trial presented in this section describe those that were detailed in the ARCH Clinical Study Report (CSR) and were determined using the standard methodology of last observation carried forward (LOCF) imputation for missing data, as pre-specified in the statistical analysis plan. However, the data more recently presented in the peer-reviewed New England Journal of Medicine publication regarding fractures and bone mineral density (BMD) were determined using a multiple imputation for the missing data as requested by the journal, which does not reflect the original pre-specified analyses, and has thus not been included in this submission.” Please clarify whether there were any differences in estimates of effect between the two methods of imputation, and describe how any differences between these analyses could affect the cost-effectiveness estimate.

The Statistical Analysis Plan (SAP) pre-specified the LOCF to assess the fracture efficacy. The New England journal requested an alternative assessment using multiple imputation. The methodology used to derive the clinical effectiveness for vertebral fractures in the ARCH trial had no bearing on the results. The New England publication provides the results for the two methods of imputation results in the Supplementary Materials (Table S1). These are reproduced below:

- HR for New Vertebral Fractures at 12 months were 0.63 (0.47-0.85) and 0.64 (0.46-0.89) using Multiple Imputation and LOCF, respectively.
- HR for New Vertebral Fractures at 24 months were 0.52 (0.40-0.66) and 0.50 (0.38-0.66) using Multiple Imputation and LOCF, respectively.

Full results with both methods of imputation are presented in Table 2, taken from the Supplementary Appendix of the publication.[30]

Table 2: Fracture endpoints at pre-specified timepoints

	Alendronate to Alendronate (N = 2047) % (n/N1)	Romosozumab to Alendronate (N = 2046) % (n/N1)	Risk Ratio or Hazard Ratio (95% CI)	Nominal *P* Value
12 Month Double-Blind Period				
New vertebral fracture by multiple imputation[a]	6.3% (128/2047)	4.0% (82/2046)	0.63 (0.47, 0.85)	0.003
New vertebral fracture by LOCF[b]	5.0% (85/1703)	3.2% (55/1696)	0.64 (0.46, 0.89)	0.008
New or worsening vertebral fracture[b]	5.9% (101/1703)	4.0% (67/1696)	0.66 (0.49, 0.89)	0.006
Clinical vertebral fracture[c]	0.9% (18/2047)	0.5% (10/2046)	0.56 (0.26, 1.22)	0.14
Clinical fracture[c]	5.4% (110/2047)	3.9% (79/2046)	0.72 (0.54, 0.96)	0.027
Nonvertebral fracture[c,d]	4.6% (95/2047)	3.4% (70/2046)	0.74 (0.54, 1.01)	0.057
Major nonvertebral fracture[c,e]	4.3% (88/2047)	2.9% (59/2046)	0.67 (0.48, 0.94)	0.019
Hip fracture[c]	1.1% (22/2047)	0.7% (14/2046)	0.64 (0.33, 1.26)	0.19
Osteoporotic fracture[c,f]	9.2% (189/2047)	6.5% (134/2046)	0.71 (0.57, 0.88)	0.002
Major osteoporotic fracture[c,g]	4.2% (85/2047)	3.0% (61/2046)	0.72 (0.52, 1.01)	0.053
Month 24				
New vertebral fracture by multiple imputation[a]	11.9% (243/2047)	6.2% (127/2046)	0.52 (0.40, 0.66)	<0.001
New vertebral fracture by LOCF[b,h]	8.0% (147/1834)	4.1% (74/1825)	0.50 (0.38, 0.66)	<0.001
New or worsening vertebral fracture[b]	9.2% (168/1834)	4.8% (87/1825)	0.52 (0.40, 0.66)	<0.001
Clinical vertebral fracture[c]	2.1% (44/2047)	0.9% (18/2046)	0.41 (0.24, 0.71)	<0.001
Primary Analysis Period				
Clinical fracture[c,h]	13.0% (266/2047)	9.7% (198/2046)	0.73 (0.61, 0.88)	<0.001
Nonvertebral fracture[c,d]	10.6% (217/2047)	8.7% (178/2046)	0.81 (0.66, 0.99)	0.037
Major nonvertebral fracture[c,e]	9.6% (196/2047)	7.1% (146/2046)	0.73 (0.59, 0.90)	0.004
Hip fracture[c]	3.2% (66/2047)	2.0% (41/2046)	0.62 (0.42, 0.92)	0.015
Osteoporotic fracture[c,f]	19.1% (392/2047)	13.0% (266/2046)	0.65 (0.56, 0.76)	<0.001
Major osteoporotic fracture[c,g]	10.2% (209/2047)	7.1% (146/2046)	0.68 (0.55, 0.84)	<0.001

A14. Regarding data extraction and quality assessment **(**Appendix D, Section D.2 of the CS), please clarify how disagreements about data extraction were resolved and please also clarify which version of the Cochrane Risk of Bias tool for RCTs was used.

Any discrepancies between two independent blinded reviewers were resolved through mutual discussion and consensus, and if not achieved, a third independent reviewer was involved to justify correct choices on extracted data and RCT quality.

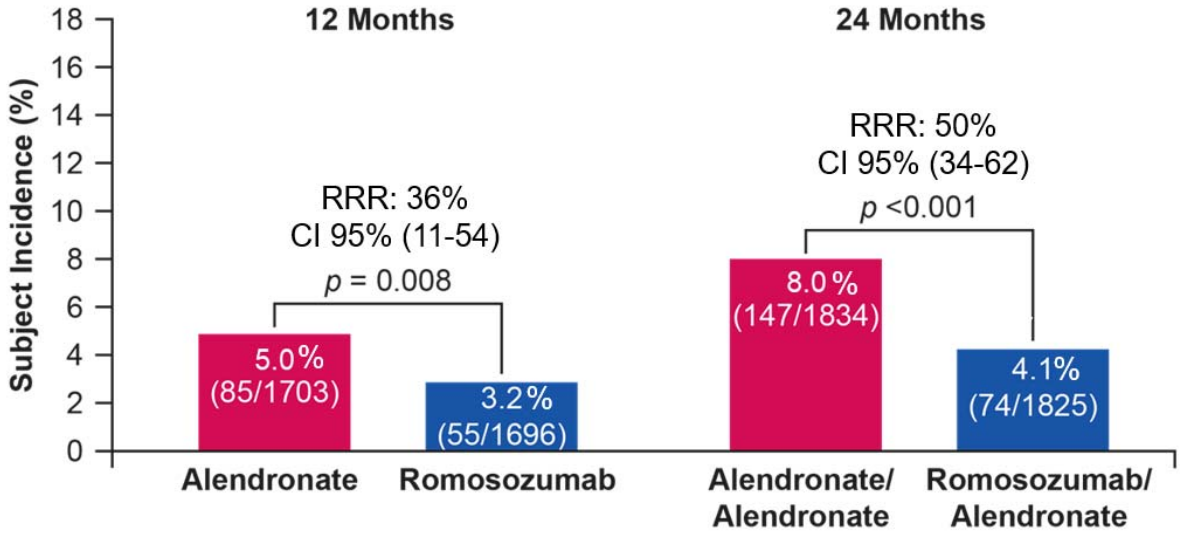
The risk of bias was assessed using the Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0.[31]

A15. Please include 95% confidence intervals when reporting any and all effect estimates, relevant throughout, but particularly in Section B.2.6 of the CS. Also, please include 95% confidence intervals on bar charts presented in Figures 5, 6, 8 in

Document B. Please also add the number of patients to Figure 8 in Document B, as in Figures 5 and 6 in Document B.

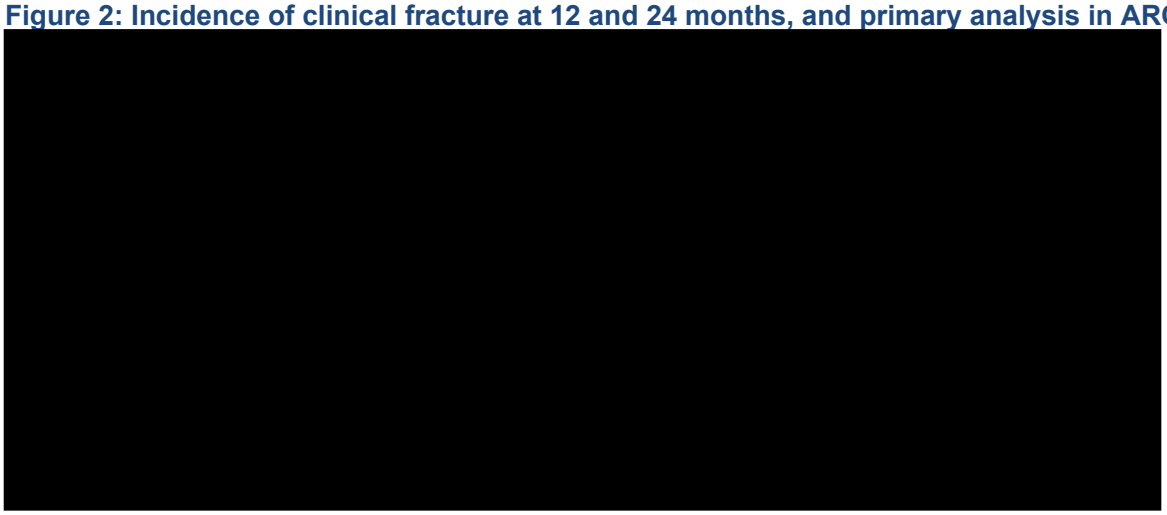
Updated versions of Figure 5, 6 and 8 (in Document B) including the requested information are presented in Figure 1, Figure 2 and Figure 3 below.

Figure 1: Incidence of new vertebral fracture at 12 and 24 months in ARCH[a]



Footnotes: [a] Number of patients in each arm is the number of subjects in the primary analysis set for vertebral fractures.
Abbreviations: RRR: relative risk reduction.
Source: Adapted from ARCH clinical study report.[32]

Figure 2: Incidence of clinical fracture at 12 and 24 months, and primary analysis in ARCH



Abbreviations: RRR: relative risk reduction.
Source: Adapted from ARCH clinical study report.[32]

Figure 3: Incidence of non-vertebral, major osteoporotic and hip fractures at primary analysis

Footnotes: [a] Adjusted 2-sided p value presented for incidence of non-vertebral fractures
Abbreviations: RRR, relative risk reduction.
Source: Adapted from ARCH clinical study report.[32]

A16. Please further justify why the ARCH intention-to-treat (ITT) population is generalisable to the UK treatment population, including how many patients in ARCH were from the UK, and whether the demographics of the ARCH ITT population match that of the UK treatment population (particularly ethnic group and geographic region), and if not, how any differences are likely to affect the cost-effectiveness estimate.

The Phase 3 study ARCH was a multicentre international study which included 533 patients (13%) from Western Europe, Australia and New Zealand. From the UK, 76 patients (1.9%) were enrolled in the trial. No regional differences were seen in the efficacy results of the ARCH trial and therefore it is reasonable to say that this will be representative of the UK population.[32] In addition, in a UK simulation of women aged 50 years and older who would typically be assessed for fracture risk, 10% of patients categorised as very high risk of fracture were representative of those enrolled in the Phase 3 ARCH study.[11]

A17. Please clarify whether in Table 6 of Document B "A 25-hydroxyvitamin D level of >20 ng/ml" should be "A 25-hydroxyvitamin D level of <20 ng/ml", given the median and 25th centile (in Table 7 of Document B) are both above 20 ng/ml for the ARCH trial population.

UCB can confirm that Table 6 should read "A 25-hydroxyvitamin D level of <20 ng/ml" as suggested by the ERG, instead of "A 25-hydroxyvitamin D level of >20 ng/ml".

A18. Please justify why the per protocol analysis set was used for some outcomes rather than the ITT analysis set. Please also provide the ITT results for the incidence

of new vertebral fractures and clarify for all analyses which analysis set is being used.

As detailed in Document B, Section B.2.4.1, Table 3 below details the different patient populations in the ARCH trial, the outcomes that each population was used to analyse, and the justification for the use of each analysis set.

Table 3: Trial populations for ARCH

Analysis	NCT01631214 (ARCH)	Outcomes assessed	Justification
Full analysis set	• Included all randomised patients in the trial. They were analysed according to their randomised treatment assignments	• Nonvertebral fracture • Clinical fracture • Clinical vertebral fracture • All fracture • Major nonvertebral fracture • Major osteoporotic fracture • Hip fracture.	This analysis set was used for the ITT analyses
Primary efficacy analysis set	• Included all randomised patients who had a baseline and ≥1 post-baseline evaluation of vertebral fracture at or before the timepoint of consideration • Patients were analysed according to their randomised treatment assignments • Patients whose first post-baseline spinal radiograph showed no fracture on vertebra, but who had the same vertebrae at baseline were also included as it could be inferred that their baseline scores would have also reported no fracture, had they been available • In this set there are more than 80% of patients from the full analysis set: o 24 Month: Alendronate 1834/2047 = 89.6% Romosozumab 1825/2046 = 89.2%	• New vertebral fractures • New or worsening vertebral fractures • Multiple new or worsening vertebral fractures	To assess new or worsening vertebral fractures, comparisons of baseline and a later assessment were necessary

Analysis	NCT01631214 (ARCH)	Outcomes assessed	Justification
	o 12 Month: Alendronate 1703/2047 = 83.2% Romosozumab 1696/2046 = 82.9%		
Per protocol analysis set	• Included patients in the full analysis set (for clinical and non-vertebral fracture) and the primary efficacy analysis set for vertebral fractures (for new vertebral fractures) who received active investigational products and met all of the patient eligibility criteria	• Clinical fracture • New vertebral fracture, and nonvertebral fracture through month 24 • Clinical fracture and nonvertebral fracture at primary analysis • Nonvertebral fracture at final analysis	This analysis set was used for sensitivity analyses only
Safety analysis set	• Patients who received ≥1 active dose of investigational product in the 12-month double-blind study period were included in this study set	• Safety data analysis for the double-blind study period, primary analysis period, and overall study period used this safety analysis set	N/A

Abbreviations: ITT: intention-to-treat; MOF: major osteoporotic fracture
Sources: ARCH Clinical Study Report.[32]

Indirect comparisons

A19. For all network meta-analysis (NMA) closed-loop analyses, please provide both the fixed-effect and random-effects estimates for the direct and indirect effects to placebo (for all treatments in the loop) so we can assess whether the inconsistency factors show whether the lack of statistically significant inconsistency is due to a lack of statistical power.

ITT Population

New vertebral fractures

a. New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Table 4: New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) - fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo		[redacted]	[redacted]

Alendronate			
Romosozumab			

Table 5: New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) - random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Fixed effects results

Table 6: New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Table 7: New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

c. New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 8: New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			

Teriparatide			
Abaloparatide			

Table 9: New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Non-vertebral fractures

a. Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Fixed effects

Table 10: Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

Table 11: Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 12: Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Random effects

Table 13: Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

c. Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 14: Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Table 15: Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

d. Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 16: Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Table 17: Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Hip fractures

a. Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Table 18: Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

Table 19: Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 20: Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Table 21: Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

EU LABEL population

New vertebral fractures

a. New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Table 22: New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

Table 23: New vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 24: New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Table 25: New vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

c. New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 26: New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Table 27: New vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Non-vertebral fractures

a. Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Table 28: Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

Table 29: Non-vertebral fractures – 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 30: Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Table 31: Non-vertebral fractures – 12 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			

Teriparatide			

c. Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 32: Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Table 33: Non-vertebral fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

d. Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 34: Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Risedronate	Teriparatide
Placebo			
Risedronate			
Teriparatide			

Table 35: Non-vertebral fractures – 24 months (Risedronate, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			

Teriparatide			
Abaloparatide			

Hip fractures

a. Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI)

Table 36: Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

Table 37: Hip fractures - 12 months (Alendronate, Romosozumab, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Alendronate	Romosozumab
Placebo			
Alendronate			
Romosozumab			

b. Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI)

Table 38: Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – fixed effects results

	Placebo	Teriparatide	Abaloparatide
Placebo			
Teriparatide			
Abaloparatide			

Table 39: Hip fractures – 24 months (Abaloparatide, Teriparatide, Placebo), Relative Risk (95%CrI) – random effects results

	Placebo	Teriparatide	Abaloparatide

Placebo			
Teriparatide			
Abaloparatide			

A20. Table 41 in Appendix D (page 143) is missing data for FRAME and Chao 2013, and Figure 16 in Appendix D is missing the Hadji 2012 study – please check all tables and figures in sections D.4.3, D.4.4 and D.4.5 to ensure all studies are included in the figures and in the tables and vice versa. Please also add percentages to the Events/N columns for all tables in these sections.

Please see Appendix A20 of the appendices to the clarification questions for the updated tables.

A21. PRIORITY QUESTION: The NMAs for BMD outcomes used the final time points of all included studies, unlike in the fracture outcomes where NMAs were specific to different time-points.

- **Please justify why the latest time-points were used, rather than splitting the NMAs into separate time points. If due to a lack of data, please justify this by showing the networks of evidence that would be for studies with outcomes at 12, 24 and 36 months.**

- **If feasible, please conduct separate NMAs using studies with outcomes at 12, 24 and 36 months for all BMD outcomes. If this is not feasible, please explain why and describe what effect combining different time-points may have on the results of the BMD NMAs.**

- **Please also add the time-point of analysis to all studies for all tables in Appendix D.4.5.**

BMD endpoints in the NMA were not presented separately by timepoint due to a paucity of information of time-specific changes in BMD outcomes in RCTs identified in the SLR, as can be observed from the tables and networks of evidence presented below.

The BMD data availability presented in the tables below results in networks of evidence that become smaller over time, which limits time point specific comparison for BMD endpoints.

Table 40: Total hip BMD data

RCT	Arm 1	Arm 2	Arm 3	Arm 4	Time point data available per study			Time point used in present BMD analyses
					12 months	24 months	36 months	
ACTIVE	Teriparatide	Placebo	Abaloparatide					18 months
Neer et al.	Teriparatide	Placebo						21 months
DEFEND	Denosumab	Placebo						24 months
SPIMOS	Ibandronate	Placebo						12 months
McClung et al. 2009	Ibandronate	Placebo						12 months
NCT00132808	Zoledronate	Placebo						24 months
FOSIT	Alendronate	Placebo						12 months
Adami et al. 1995	Alendronate	Placebo						24 months
Tucci et al.	Alendronate	Placebo						36 months
Silverman et al.	Raloxifene	Placebo						36 months
FRAME	Romosozumab	Placebo						36 months
Hadji et al. 2012	Teriparatide	Risedronate						18 months
DATA	Denosumab	Teriparatide						24 months
STRUCTURE	Romosozumab	Teriparatide						12 months
Recknor et al.	Ibandronate	Denosumab						12 months
Miller et al. 2016	Zoledronate	Denosumab						12 months
DECIDE	Alendronate	Denosumab						12 months
STAND	Alendronate	Denosumab						12 months
Tan et al.	Alendronate	Zoledronate						36 months
ARCH	Romosozumab	Alendronate						36 months
EUROFORS	Placebo	Teriparatide	Raloxifene					24 months
Amgen 20010223	Placebo	Denosumab	Alendronate					48 months
FACTS1	Placebo	Risedronate	Alendronate					24 months
McClung et al. 2014	Placebo	Teriparatide	Alendronate	Romosozumab				12 months
HORIZON	Zoledronate	Placebo						36 months
Grey et al.	Zoledronate	Placebo						12 months
Roux et al. 2013	Denosumab	Risedronate						12 months
FACT	Alendronate	Risedronate						12 months
MOTION	Alendronate	Ibandronate						12 months
EFFECT international	Alendronate	Raloxifene						12 months

Networks of evidence for total hip BMD

Figure 4: 12 months (10 treatments, 26 RCTs)

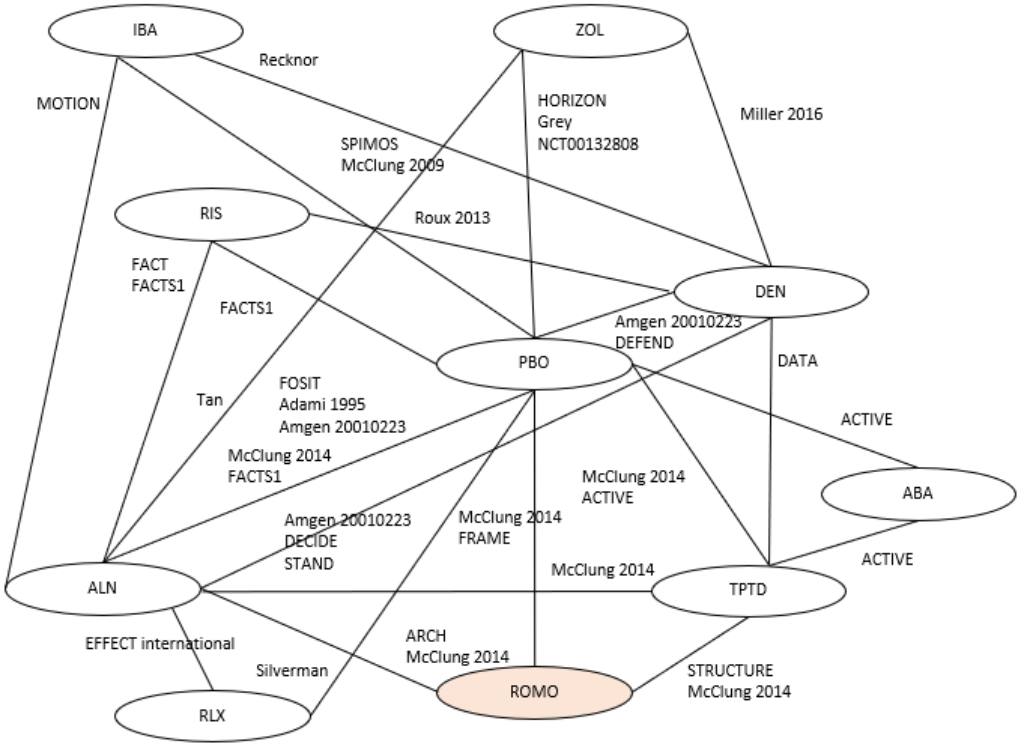


Figure 5: 24 months (8 treatments, 13 RCTs)

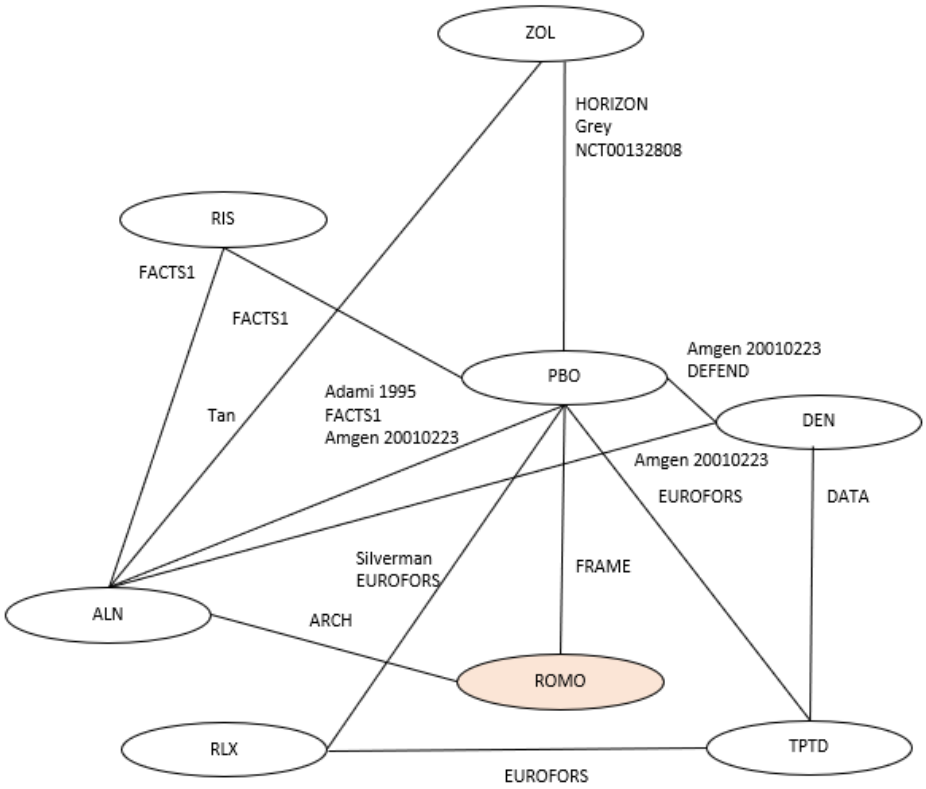


Figure 6: 36 months (6 treatments, 8 RCTs)

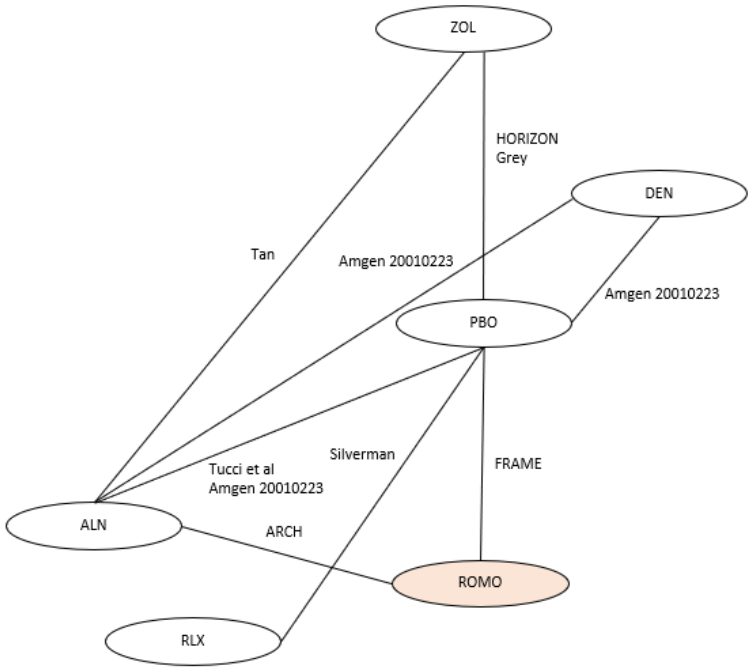


Table 41: Femoral Neck BMD data

RCT	Arm 1	Arm 2	Arm 3	Arm 4	Time point data available per study			Time point used in present BMD analyses
					12 months	24 months	36 months	
Fogelman et al.	Risedronate	Placebo						24 months
ACTIVE	Teriparatide	Placebo	Abaloparatide					18 months
Neer et al.	Teriparatide	Placebo						21 months
DEFEND	Denosumab	Placebo						24 months
McClung et al. 2009	Ibandronate	Placebo						12 months
NCT00132808	Zoledronate	Placebo						24 months
Dursun et al.	Alendronate	Placebo						12 months
FOSIT	Alendronate	Placebo						12 months
Adami et al. 1995	Alendronate	Placebo						24 months
Aki et al.	Alendronate	Placebo						12 months
Tucci et al.	Alendronate	Placebo						36 months
NCT00398606	Alendronate	Placebo						24 months
Adami et al. 2008	Raloxifene	Placebo						24 months
FRAME	Romosozumab	Placebo						36 months
Hadji et al. 2012	Teriparatide	Risedronate						18 months
DATA	Denosumab	Teriparatide						24 months
STRUCTURE	Romosozumab	Teriparatide						12 months
Recknor et al.	Ibandronate	Denosumab						12 months
DECIDE	Alendronate	Denosumab						12 months
Tan et al.	Alendronate	Zoledronate						36 months
ARCH	Romosozumab	Alendronate						36 months
EUROFORS	Placebo	Teriparatide	Raloxifene					24 months
Um et al. 2017	Placebo	Alendronate	Raloxifene					36 months
Johnell et al.	Placebo	Alendronate	Raloxifene					12 months
FACTS1	Placebo	Risedronate	Alendronate					24 months
McClung et al. 2014	Placebo	Teriparatide	Alendronate					12 months
HORIZON	Zoledronate	Placebo						36 months
Liberman et al.	Alendronate	Placebo						36 months
Roux et al. 2013	Denosumab	Risedronate						12 months
FACT	Alendronate	Risedronate						12 months
EFFECT international	Alendronate	Raloxifene						12 months

Networks of evidence for **Femoral Neck BMD**

Figure 7: 12 months (10 treatments, 24 RCTs)

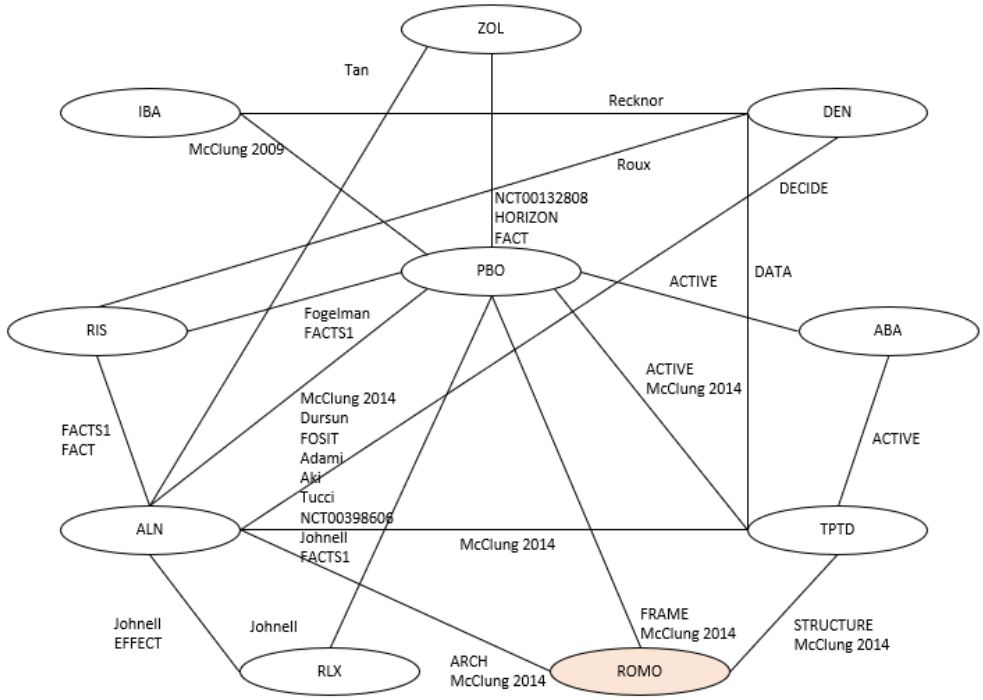


Figure 8: 24 months (8 treatments, 14 RCTs)

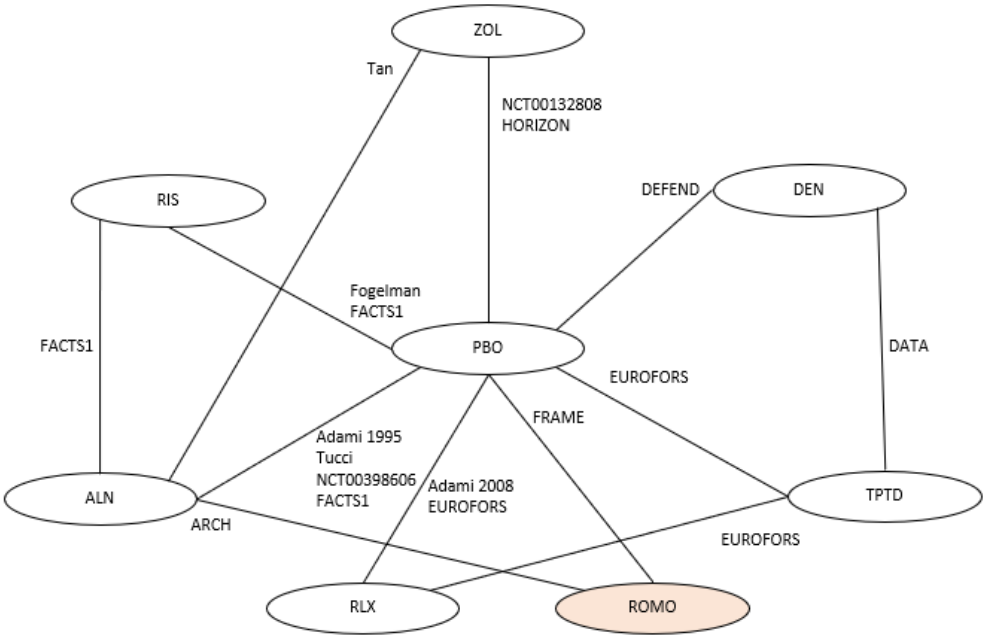


Figure 9: 36 months (5 treatments, 7 RCTs)

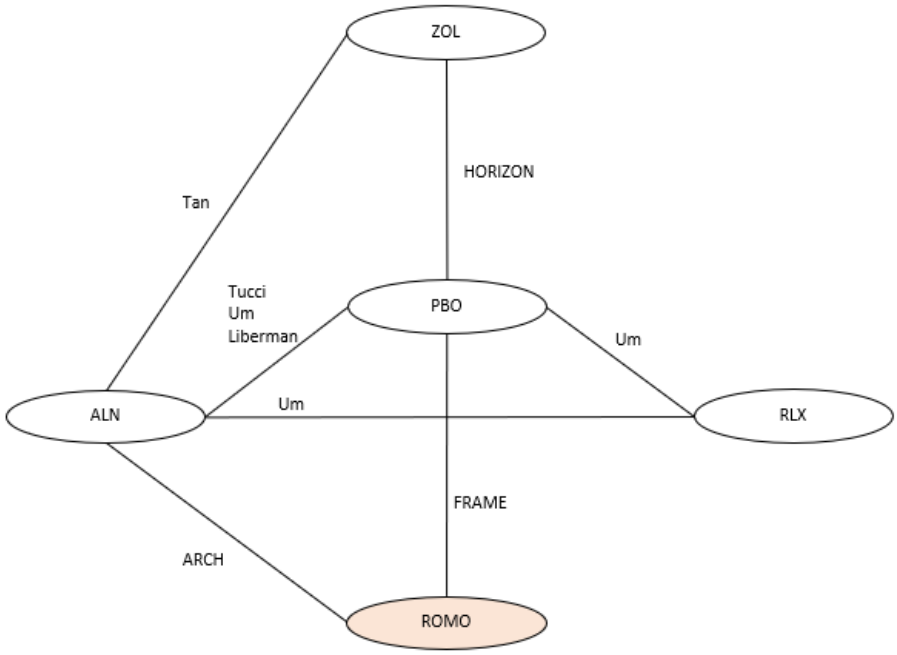


Table 42: Lumbar Spine BMD data

RCT	Arm 1	Arm 2	Arm 3	Arm 4	Time point data available per study			Time point used in present BMD analyses
					12 months	24 months	36 months	
NCT00353080	Risedronate	Placebo						24 months
Fogelman et al.	Risedronate	Placebo						24 months
ACTIVE	Teriparatide	Placebo	Abaloparatide					18 months
Neer et al.	Teriparatide	Placebo						21 months
DEFEND	Denosumab	Placebo						24 months
SPIMOS	Ibandronate	Placebo						12 months
NCT00132808	Zoledronate	Placebo						24 months
Dursun et al.	Alendronate	Placebo						12 months
FOSIT	Alendronate	Placebo						12 months
Adami et al. 1995	Alendronate	Placebo						24 months
Aki et al.	Alendronate	Placebo						12 months
Tucci et al.	Alendronate	Placebo						36 months
NCT00398606	Alendronate	Placebo						24 months
OCEAN	Alendronate	Placebo						12 months
Adami et al. 2008	Raloxifene	Placebo						24 months
Silverman et al.	Raloxifene	Placebo						36 months
FRAME	Romosozumab	Placebo						36 months
Hadji et al. 2012	Teriparatide	Risedronate						18 months
DATA	Denosumab	Teriparatide						24 months
STRUCTURE	Romosozumab	Teriparatide						12 months
Recknor et al.	Ibandronate	Denosumab						12 months
Miller et al. 2016	Zoledronate	Denosumab						12 months
DECIDE	Alendronate	Denosumab						12 months
STAND	Alendronate	Denosumab						12 months
Tan et al.	Alendronate	Zoledronate						36 months
ARCH	Romosozumab	Alendronate						36 months
EUROFORS	Placebo	Teriparatide	Raloxifene					24 months
Amgen 20010223	Placebo	Denosumab	Alendronate					48 months
Um et al. 2017	Placebo	Alendronate	Raloxifene					36 months
Johnell et al.	Placebo	Alendronate	Raloxifene					12 months
FACTS1	Placebo	Risedronate	Alendronate					24 months
McClung et al. 2014	Placebo	Teriparatide	Alendronate	Romosozumab				12months
HORIZON	Zoledronate	Placebo						36months
Grey et al.	Zoledronate	Placebo						12months

RCT	Arm 1	Arm 2	Arm 3	Arm 4	Time point data available per study			Time point used in present BMD analyses
					12 months	24 months	36 months	
Reid et al.	Zoledronate	Placebo						6 years
Liberman et al.	Alendronate	Placebo						36 months
Roux et al. 2013	Denosumab	Risedronate						12months
FACT	Alendronate	Risedronate						12months
MOTION	Alendronate	Ibandronate						12months
EFFECT	Alendronate	Raloxifene						12months

Networks of evidence for Lumbar Spine BMD

Figure 10: 12 months (10 treatments, 34 RCTs)

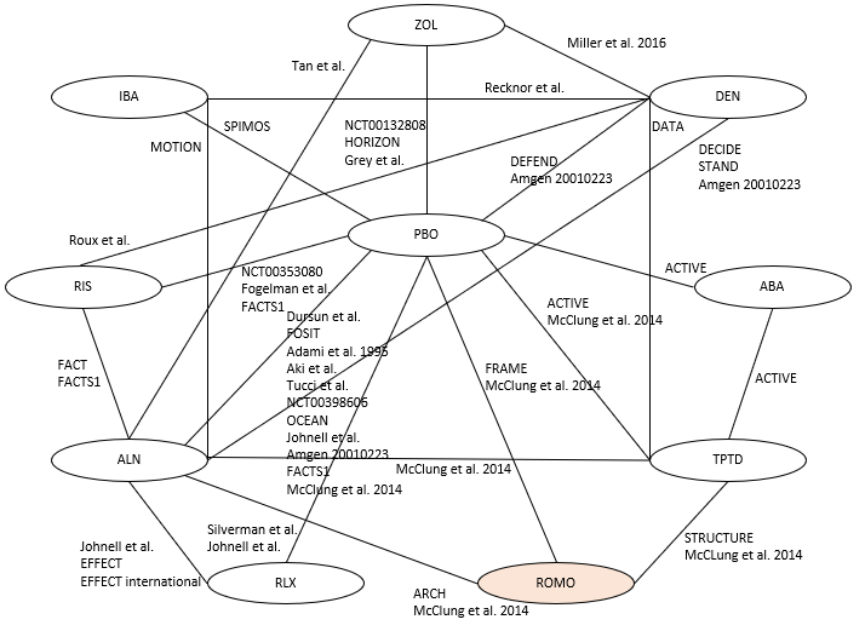


Figure 11: 24 months (10 treatments, 34 RCTs)

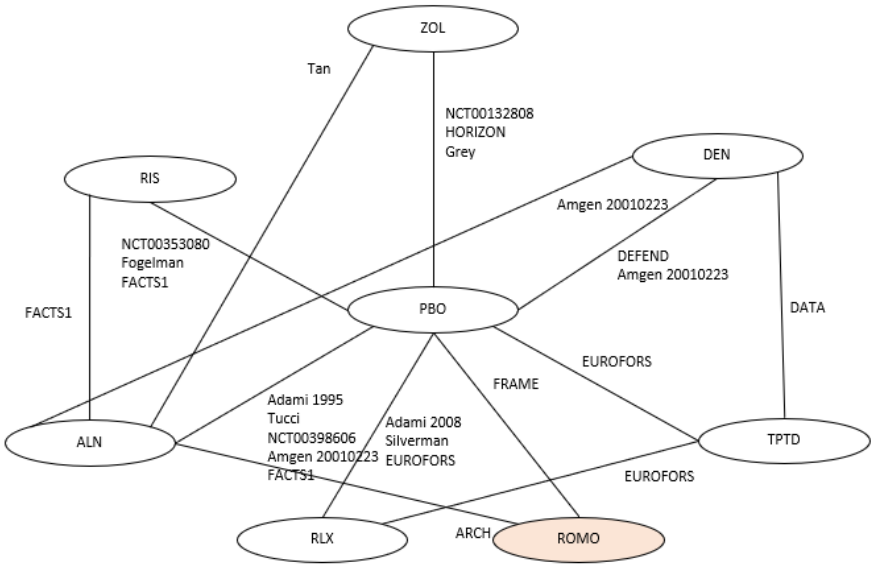
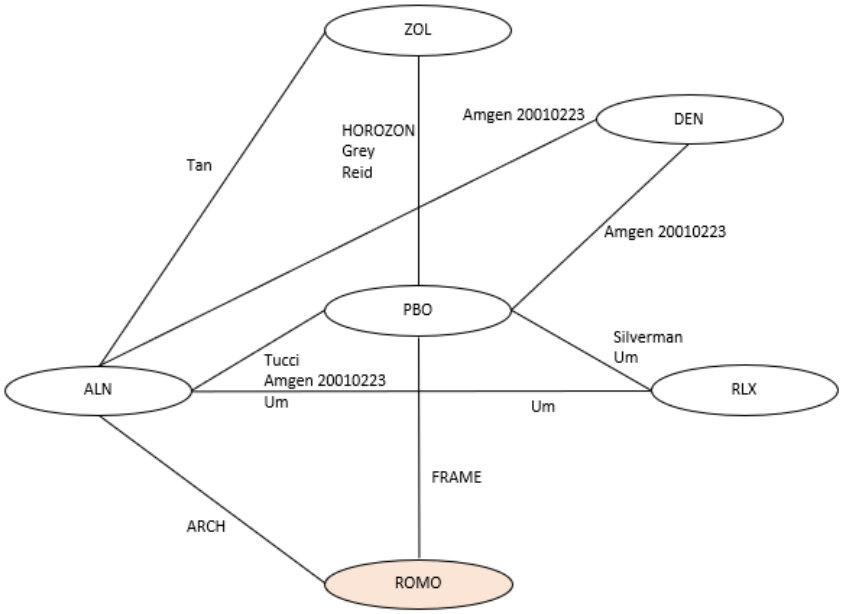


Figure 12: 36 months (6 treatments, 10 RCTs)



Based on the above data situation, it was not considered appropriate to conduct separate NMAs with outcomes at 12, 24 and 36 months for all BMD outcomes.

Combining BMD outcomes across time-points can potentially be considered “at risk” of neglecting differences in onset of action, but nevertheless, this approach has been adopted in previous NMAs by NICE.[33, 34]

Whilst acknowledging the limitations of the BMD NMA results, it should be noted that the cost-effectiveness model for romosozumab does not consider any BMD outcomes, and so these limitations do not impact the cost-effectiveness results of any technologies. The exclusion of BMD outcomes could be considered conservative for romosozumab, because it has demonstrated superior gains in BMD against alendronate, placebo/denosumab and teriparatide in ARCH, FRAME and STRUCTURE, respectively, as shown in the CS Document B, Figure 10 and Figure 11 (ARCH), and the CS Appendices, Figure 78 (FRAME) and Figure 79 (STRUCTURE).

A22. PRIORITY QUESTION: Please provide all analysis code for all analyses, including the WinBugs code and input data for the NMAs.

Please see Appendix A22 of the appendices to the clarification questions for the analysis code for all analyses. UCB remain available to address any further queries related to the use of the provided codes and input data.

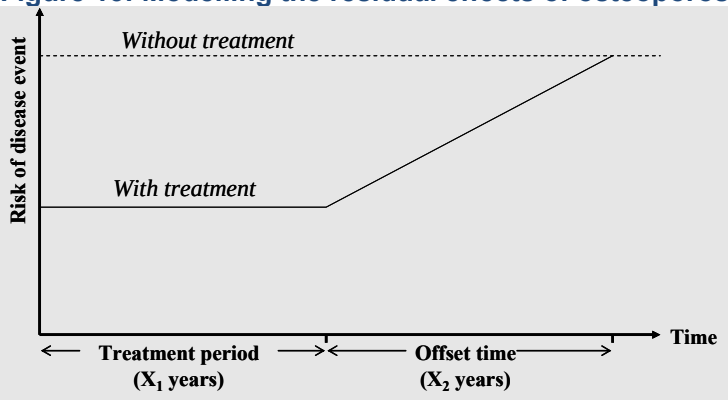
Section B: Clarification on cost-effectiveness data

Model structure and implementation

B1. PRIORITY QUESTION: Please define *all treatment sequences* included in the cost effectiveness analyses. This should include the base-case, the scenario analyses and the complete time horizon, *indicating also what effects are maintained and for how long*. This could be presented in the form of a table as below (please add rows/columns if needed):

As detailed in Document B, Section B.3.3.5, dynamic residual effects are applied to almost all of the treatment sequences including in the base case and scenario analysis comparisons. The economic model assumed that the offset time associated with each treatment was equal to the time a patient remained on treatment – during this offset time, the fracture risk reduction was assumed to decline linearly to zero – an example is presented in Figure 13 below.

Figure 13: Modelling the residual effects of osteoporosis treatments



Abbreviations: X1: treatment period; X2: offset time of treatment effect.

The only exception is denosumab – for denosumab, the clinical effect is reported to be limited to within six months after stopping treatment.[35, 36] As such, a conservative one-year fixed offset time is applied to denosumab in the economic model.

A summary of the treatment sequences and associated length of effects are detailed in Table 43 below, and further details about how dynamic residual effects are applied within the model are presented in response to Question B.11.

Table 43: Summary of the treatment sequences and effects applied for base case and scenario comparisons

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11+
Base case comparisons											
Intervention: Romosozumab/ alendronate	ROMO	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Base case comparison 1: Alendronate	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenarios											
Scenario 1: Alendronate	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 2: Teriparatide (Forsteo) 24 months	TRP	TRP	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect
Scenario 3: Teriparatide (Forsteo) 18 months	TRP	TRP (6 months) NONE (6 months)	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect

Scenario 4: Teriparatide (biosimilar Movymia) to alendronate	TRP	TRP (6 months) ALN (6 months)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 5: Teriparatide (Forsteo) to alendronate	TRP	TRP (6 months) ALN (6 months)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 6: Raloxifene	RAL	RAL	RAL	RAL	RAL	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 7: Denosumab	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect
Scenario 8: Risedronate	RIS	RIS	RIS	RIS	RIS	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 9: Zoledronate	ZOL	ZOL	ZOL	ZOL	ZOL	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 10: Alendronate	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE

Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 11: Denosumab	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Treatment-dependent fracture risk reduction	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect

Abbreviations: ALN: alendronate; DEN: denosumab; RAL: raloxifene; ROMO: romosozumab; TRP: teriparatide; ZOL: zoledronate.

B2. On page 68 of the CS, it is mentioned that "All patients started the model in the "at risk" health state. At the end of each cycle patients either *moved* into the one of the fracture states, remained in their current health state without new fracture, or died." Please explain how transitions are determined in the model (e.g., by a random draw from different probability distributions).

The model is created as a dynamic population microsimulation model that tracks every individual within the model and stores all necessary information on an individual level.

The underlying decision processes in the model e.g., mortality, fracture events, treatment discontinuation, are completed using a similar set of rules and calculations. For every individual processed through the simulation, a uniformly distributed random number is used for determining the next state for an individual. The random number is generated using the Mersenne twister algorithm[37] for pseudo-random number generation. While many other random number generators are available, the Mersenne Twister generator was chosen since it is computationally fast, easily implemented in VBA and has a sufficiently long period length (219937-1, i.e., the number of steps before the program starts repeating itself) for disease simulation models.

A random seed is used to set a starting point for generating a series of random numbers and thereby produce the identical results each time the model is run with the exact same settings. Without random seed, the results will always fluctuate slightly, although less and less as the number of iterations increase. The user may choose to use/not to use the random seed (sheet "Misc", cell C46).

A standard technique which was implemented in the model to reduce the stochastic noise between model runs is synchronised random numbers (or "common random numbers").[38] With synchronised random numbers, the same random number sequences are used within simulated individuals across comparators. When used in addition to random seeding, the model generates identical individuals across each model run and thus each individual can serve as her own control for counterfactual analysis. Synchronised random numbers do not, however, reduce variation within a single model run which simulating a large number of iterations might. The remaining variation primarily arises from changing model parameters or assumptions.

Thereafter, the data of interest is loaded containing probabilities of an event of interest. The probability is compared with the random number generated using formula 1:

$$Decision = \begin{cases} Random\ number \geq probability, 1 \\ Random\ number < probability, 0 \end{cases} \quad \text{(formula 1)}$$

Where the probability is not a single number, but instead a list of probabilities, a calculation is used for converting the list to cumulative probabilities as:

1. Original probability list for four possible outcomes: [0.4,0.2,0.1,0,3]
2. The list is recalculated for every element to become a cumulative sum of the previous elements, i.e.;
3. Recalculated probability list for the four possible outcomes: [0.4,0.6,0.7,1]
4. The random number generated within the model is compared to every element returning 1 for the element fulfilling:

$$Decision = \begin{cases} Random\ number \geq probability_i, 1 \\ Random\ number < probability_i, 0 \end{cases}$$ (formula 2)

Hence, the generated random number is compared with the list elements and returns the state with a probability closest, however larger than the random number, as illustrated in Formula 2.

B3. On page 66 of the CS, it is mentioned that "The algorithm used to generate the estimated fracture risk within the model is based on FRAX, but also includes an additional risk associated with recent fracture". Please provide a numerical example illustrating how FRAX and the additional risk associated with recent fracture are used in the economic model. Please clarify whether this additional risk has been validated by experts. Finally, please conduct a scenario analysis based on FRAX only.

While it is well established that a fragility fracture increases the risk of a subsequent fracture over a patient's lifetime, recent studies have shown that the increase in relative risk may not be constant over time, age and the number of fractures.[39-41] In a review of data on identification and treatment of patients with osteoporosis at increased risk of fracture, a working group convened by the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis, concluded that it is evident that the risk of fracture is highest immediately after a fracture.[42]

FRAX accommodates the well-established risk factor of prior fragility fracture. Kanis et al. (who have developed the FRAX tool) acknowledge imminent fracture risk and describe that FRAX cannot currently accommodate recency of fracture.[43] For example, they write the following (quote):

"The immediate risk is high and then wanes over time for approximately 2 years. Thereafter, a nadir is reached but the risk remains higher than that of the general population. The early phase of particularly high risk has been termed imminent risk (…). This transiency, which is not currently accommodated in the FRAX algorithm, suggests that treatment given to such patients immediately after fracture might avoid a higher number of new fractures compared with treatment given at a later date. This reinforces a rationale for very early intervention immediately after fractures to avoid recurrent fractures. Furthermore, it mandates the use of the most effective therapies early in the course of treatment, rather than delaying their use to a time of lower fracture risk. Thus, the quantification of imminent risk enables the targeting of anabolic treatments to individuals identified to be at very high risk (…)."[43]

Thus, Kanis et al. describe the importance of adjusting FRAX to accommodate imminent risk, to correctly quantify fracture risk in a patient population who are at high need of effective and rapid treatment.[43] The currently available official FRAX algorithm does not include imminent risk. Our model adjusts the FRAX risk for imminent risk, based on data from a Swedish retrospective real-world data study. This adjustment has been validated and was accepted by clinical experts in an internal economic advisory-board that was held in 2017, in the validation process with PRIMA in 2017. The incorporation of imminent risk in the CE model has also been described in two published peer-reviewed manuscripts which provides validation of the approach.[44, 45]

The Swedish real-world data study identified a high imminent risk of subsequent major osteoporotic fracture in women with one, two and three fractures. The study is based on Swedish national register data on all individuals who were dispensed an osteoporosis drug, had a fracture, and/or had a DXA scan at one of the participating clinics from year 2000. The study is described in Söreskog et al.[46]

The model is populated with the relative risk (RR) of fracture compared with individuals without fracture, after 1st, 2nd and 3rd fracture by fracture site and age group. These numbers also are based on the Swedish retrospective study.[46] The model updates the relative risk each time a fracture is sustained. This replaces the risk contribution that FRAX provides for prevalent fracture during the period the imminent fracture risk is higher than the prevalent fracture risk contribution from FRAX. Since population incidence, FRAX and new fracture contribute with different risk contributions, the risk of double counting is very small. Fracture risk is estimated as a function of the general population risk, the RR estimated by FRAX for a given patient profile, and the maximum of the time-dependent RR of fracture and the RR of fracture as estimated by FRAX:

$$MAX(RR_{recent\ fx\ vs\ no\ fx} \mid FRAX\ RR_{fx\ vs\ norm\ pop.})$$
$$* FRAX\ RR_{patient\ profile\ excl.fx\ CRF} FRAX\ RR_{patient\ profile\ excl.fx\ CRF}$$
$$* General\ population\ risk * Risk\ reduction\ from\ treatment$$

A numerical example on how fracture risk, and imminent risk, is calculated in the model is provided in question B4, sub-question E.

A scenario based on FRAX should not be considered relevant for decision making in this appraisal, as such a scenario does not accurately represent the romosozumab target patient population (i.e., those with a recent fracture) and the resulting increased risk of fracture experienced by these patients. Please find a more detailed description on the importance to incorporate "recency" of fracture using FRAX in response to clarification question B4-B.

However, for illustrative purposes this scenario has been tested to address the clarification question, as presented in Table 44. The ICER of romosozumab vs alendronate increases to £34,607 due to the lower fracture risk in this patient population. It is important to reiterate that this fracture risk underestimates the fracture risk that would be experienced by patients who would be eligible for treatment with romosozumab, and therefore the results of this scenario should not be considered relevant to this submission.

Table 44: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment in all patients regardless of fracture recency (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO/ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£34,607
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£12,553

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

B4. PRIORITY QUESTION: On page 69 of the CS, there are several statements that require further clarification:

A. "FRAX is not currently capable of calculating the imminent risk as the current FRAX tool does not consider recency or site of prior fracture".

Please clarify whether other tools (similar to FRAX) are capable to calculate this imminent risk.

There are no tools currently capable of calculating the imminent risk of a fragility fracture.

FRAX does not assess the recency or site of fracture and another risk assessment tool, QFracture, doesn't assess recency or BMD, and therefore both available tools are limited in their ability to accurately assess those at imminent risk of fracture.

The NOGG multidisciplinary team are aware of this limitation and are incorporating further clinical recommendations for support around imminent risk.[4, 11] Currently, the FRAX tool suggests that healthcare professionals (HCPs) should use their clinical judgement when interpreting probabilities when faced with limitations of the tool.[3] Clinical expert insights have highlighted that imminent fracture risk is most accurately assessed by FLS services following a patient recently experiencing a MOF.

B. "Therefore, the 10-year risk from FRAX will be an underestimation of the short-term fracture risk in patients who have experienced a recent fragility fracture and are at imminent risk of another fracture". Please provide an indication of the magnitude of the underestimation. Please explain how 10-year risks are converted into 6-month transition probabilities.

Using the 10-year fracture probability of a person with any fracture at any time point in the past as estimated by FRAX, underestimates the fracture risk in patients with recent fractures because the risk is highest closer to the fracture and then decreases with time. This temporal relationship has been demonstrated in several studies.[39-41, 46]

Kanis et al. (2020) published a comparison of 10-year probability of MOF for patients with a prior fracture (at any time) compared to patients with a recent fracture, based on a population-based study including Icelandic women.[47] For example, for a 70-year-old woman with a prior fracture in adult life (at any time), the 10-year probability was reported to be 27.6% (this corresponds to what FRAX estimates without imminent risk). For a 70-year-old woman with recent vertebral fracture (within the past 2 years), the probability was reported to be 41.9%. The ratio of probability between the 70-year-old woman with any prior fracture and woman with recent vertebral fracture is 1.52 and decreases with age (50 years: ratio 2.47, 60 years: ratio 1.86, 80 years: 1.24, 90 years: 1.04).

The cost-effectiveness model uses a combination of UK general population fracture incidences, adjustment for risk factors according to FRAX (using relative risks from FRAX), and adjustment of imminent fracture risk to calculate the 6-month transition probabilities. The model does not use the absolute 10-year fracture probability from FRAX to calculate the transition probabilities, however, the use of general population incidence adjusted using FRAX is a similar approach to modelling on the absolute risk from FRAX.

C. "In the model, whenever a patient sustained a fracture, their individual fracture risk was updated." Please provide a numerical example illustrating how the individual fracture risk is updated.

The fracture risk is updated such that it corresponds to the history of fracture events. For example, an average 70-year-old woman in UK has a yearly risk of hip fracture of 0.04. A 70-year-old woman with T-score -2.9 and a prior fracture (unknown site) 10 years ago, BMI 25.4, and no other risk factors, has a relative risk of 2.1 (calculated using the official FRAX algorithm). When this woman suffers a vertebral fracture, her risk of MOF in the following 6 months increases with a ratio of 4.3 (i.e., the relative risk of MOF in a patient with a vertebral fracture in the past 6 months). The increase in risk immediately after the fracture ("imminent risk") was estimated using the Swedish register data and included in the model (Söreskog et al. (2021)[45] see sheet "Recent RR input" in the model). Thus, her risk in the first 6-months after the vertebral fracture corresponds to 0.04*2.1*4.3. In the subsequent 6 months, the risk, due to the recent fracture, is increased by 2.3 instead of 4.3 (based on the Swedish register data).[33] Hence, the risk in month 7–12 after the vertebral fracture corresponds to 0.04*2.1*2.3.

D. "Although estimates of absolute fracture values vary between countries, relative estimates can be assumed to be transferable across geographic settings." Please provide evidence to support this statement.

Due to a lack of comparable studies in other countries that have estimated the relative risk of subsequent fractures, it is difficult to conclude with certainty that the relative risks would be exactly the same in Sweden and the UK.

However, it is likely that the relative risks would be similar, given the UK and Sweden are two countries which are geographically close to each other and would be expected to provide a similar quality of healthcare. For example, the European SCOPE study (which summarises key indicators of the burden of osteoporosis and its management in the EU and UK) showed that UK and Sweden have similar availability to DXA and fracture liaison services, as well as similar treatment gaps (66% in UK vs 67% in Sweden, measured as the difference in number treated for osteoporosis and number who have a fracture probability exceeding that of a woman with a prior fracture.[48] Similar assumptions, that relative estimates are comparable across countries, have been made in several studies in the published literature. For example, Hernlund et al. (2013) in a study endorsed by the International Osteoporosis Foundation, applied Swedish relative risks of death to compute absolute risk of death after fracture in other countries, such as the UK.[49]

Previous cost-effectiveness studies have made such assumptions on fracture risk and mortality after fracture when there was a lack of country-specific data.[50]

Furthermore, an independent academic Assessment Group considered it appropriate to use relative risks of subsequent fracture based on Dutch data[41] in their recent economic evaluation of non-bisphosphonates.[51] The Assessment Group did not mention potential uncertainty of using relative estimates from another country for UK.

E. Please provide a numerical example illustrating how fracture risk is estimated in the economic model. Please describe all elements in the equation on page 69 of the CS.

The risk of sustaining a fracture in the model depends on three elements: the risk for an individual in the general population incurring a fracture, the increased fracture risk associated

with osteoporosis (the relative risk) and a risk reduction, if any, attributed to a treatment. Formula on page 69 of the CS:

$$MAX(RR_{recent\ fx\ versus\ no\ fx}\ |FRAX\ RR_{fx\ vs\ norm\ pop.}) * FRAX\ RR_{\ patient\ profile\ excl.fx\ CRF} * General\ population\ risk * Risk\ reduction\ from\ treatment$$

Example:

A 74-year-old woman with T-score of −2.9, with current smoking, BMI 25.4, a recent MOF, and no other risk factors starts treatment with romosozumab. The risk of hip fracture in the first cycle for this patient is calculated in the following way:

- The risk of hip fracture in the general population for a 74-year-old woman is 0.0057 (Singer et al).
- The relative risk of hip fracture associated with having the low T-score (-2.9) in a 74-year-old smoker with BMI 25.4 is according to FRAX 1.95 (corresponds to "$FRAX\ RR_{\ patient\ profile\ excl.fx\ CRF}$" in the equation above, from page 69 of the CS). This number is calculated in the model based on the official FRAX algorithm.
- The relative risk associated with a prior fracture according to FRAX is 1.45 ("$FRAX\ RR_{fx\ vs\ norm\ pop.}$" in the equation above). This number is calculated in the model based on the official FRAX algorithm.
- The relative risk associated with a recent MOF is 2.13 ("$RR_{recent\ fx\ versus\ no\ fx}$" in the equation above, based on Swedish register data as described on page 69 in the CS).
- The risk reduction from treatment for romosozumab is 0.69 (relative risk) in the first cycle.
- Since 2.13, the relative risk of recent MOF, is higher than 1.45, the relative risk of prior fracture according to FRAX, the calculation of hip fracture risk is as follows: $MAX(RR_{recent\ fx\ versus\ no\ fx}\ |FRAX\ RR_{fx\ vs\ norm\ pop.}) * FRAX\ RR_{\ patient\ profile\ excl.fx\ CRF} * General\ population\ risk * Risk\ reduction\ from\ treatment$ = 2.13 * 1.95 * 0.0057 * 0.69
- The relative risk of recent fracture ($RR_{recent\ fx\ versus\ no\ fx}$) is used instead of relative risk of prior fracture according to FRAX ($FRAX\ RR_{fx\ vs\ norm\ pop.}$) as long as $RR_{recent\ fx\ versus\ no\ fx}$ is higher than $FRAX\ RR_{fx\ vs\ norm\ pop.}$ The relative risk of recent fracture is updated every cycle since it is time-dependent, and if it is lower than $FRAX\ RR_{fx\ vs\ norm\ pop.}$, then $FRAX\ RR_{fx\ vs\ norm\ pop.}$ would be the first part of in the formula above. With this approach, double-counting risk contribution from a fracture is avoided, since the model replaces the risk contribution of prior fracture from FRAX with the risk contribution from recent fracture (and vice versa, depending on which is highest).

Clinical parameters

B5. PRIORITY QUESTION: Please clarify the following questions regarding Table 17 of the CS:

A. Please clarify whether <u>all</u> patient characteristics are representative for UK patients. This is only indicated for age but not for the other characteristics.

A recent publication by the NOGG identified that a FRAX threshold that defines a very high risk of fractures is in alignment with the characteristics of patients enrolled in a number of Phase 3 clinical trials for anabolic osteoporosis treatments, such as ARCH. The NOGG identified that ~10% of women aged 50 years and older in the UK would be characterised at very high risk of fracture.[11]

The generalisability of the other patient characteristics listed in Table 17 is considered below:

- **Sex**: In the UK, 1 in 2 women compared to 1 in 5 men will experience a fragility fracture in a lifetime[52]
- **Prior fracture**: Of an estimated 2,527,331 of postmenopausal women with a FN T-score of −2.5 (i.e confirmed osteoporosis) in the UK, 345,197 experienced a fragility fracture in 2010 (~14%)[26]
- **Mean femoral neck T-score**: In 2010, 2,527,331 women aged 50 years and older had a FN T-score of ≤ -2.5 in the UK[26]
- **Mean BMI**: In 2019 the mean BMI for women in the UK was 27.6[53]
- **Mean 10-year MOF probability**: The 10-year probability of a MOF for a postmenopausal woman with previous fracture within the last 2 years, BMI of 25 kg/m^2 and no other risk factors according to FRAX is 30%[54] In addition, a recent publication suggests that the NOGG guideline intervention threshold for very high fracture risk (i.e., anabolic therapy considered first-line) should be 1.6x the current upper assessment threshold (10-year probability of MOF ~30%)[11]

B. Please justify the choice of 30% for the 10-year MOF probability.

The 30% MOF probability is not an input setting but rather a result of the risk factors that characterise the patient population. The target patient population for romosozumab is characterised by a 75-year-old woman with a T-score of −2.9, recent MOF and a BMI of 25.4 (to be mostly in line with the ARCH population). Additional clinical risk factors such as patients' use of tobacco, alcohol, glucocorticoids or history of rheumatoid arthritis and parental hip fracture can be accounted for in the model using FRAX, which increases the 10-year MOF probability if enabled. In the base case, patients were only enabled to be tobacco users, with all additional clinical risk factors disabled, meaning the patients were simulated to have none of the aforementioned clinical risk factors expect tobacco use, which can be considered conservative t. The characterized patient population is the expected "average" patient, based on data from the ARCH trial that included a similar patient population (mean age 74, mean T-score −2.9, and all patients had prior fractures). In reality, patients would have different risk profiles, some with higher risk which is associated with a lower ICER and some with lower risk implying a higher

ICER. Please refer to sub-question C below for a description on how different risks can be attained.

C. Please explain how sensitive the model results are to changes in patient characteristics.

Cost-effectiveness results are sensitive to fracture risk in the patient population. Romosozumab is expected to be cost-effective in a population with a 10-year fracture probability of approximately 30%. A person can attain a 30%-fracture probability due to many different combinations of risk factors, i.e., higher age, low BMD T-score, smoking, glucocorticoid use, fracture history etc. In a 74-year-old woman, 30% probability is, for example, attained from having T-score of −2.9, having a prior fracture, being a smoker, and no other risk factors. At a given age (and sex), the combination of risk factors that achieves a certain fracture probability plays a minor role in cost-effectiveness results, but higher fracture probability is associated with improved cost-effectiveness. However, age has a large impact even at a given fracture probability, due to differences in remaining lifetime expectancy.

Sensitivity analyses are included in the CS where the sensitivity to increasing and decreasing the age at which treatment is started is tested, keeping the other risk factors constant. These analyses demonstrate that the ICER is highest in the younger ages (50–60), slightly higher than the base case at age of 70 (start age in the base case is 74), and lower than the base case at age 80. It should be noted that only age is varied in these analyses which means that the fracture probability according to FRAX is lower in the age groups 50–70 compared with the base case, and higher than the base case in the age group 80.

B6. PRIORITY QUESTION: Please clarify the following questions regarding <u>baseline fracture incidence</u>:

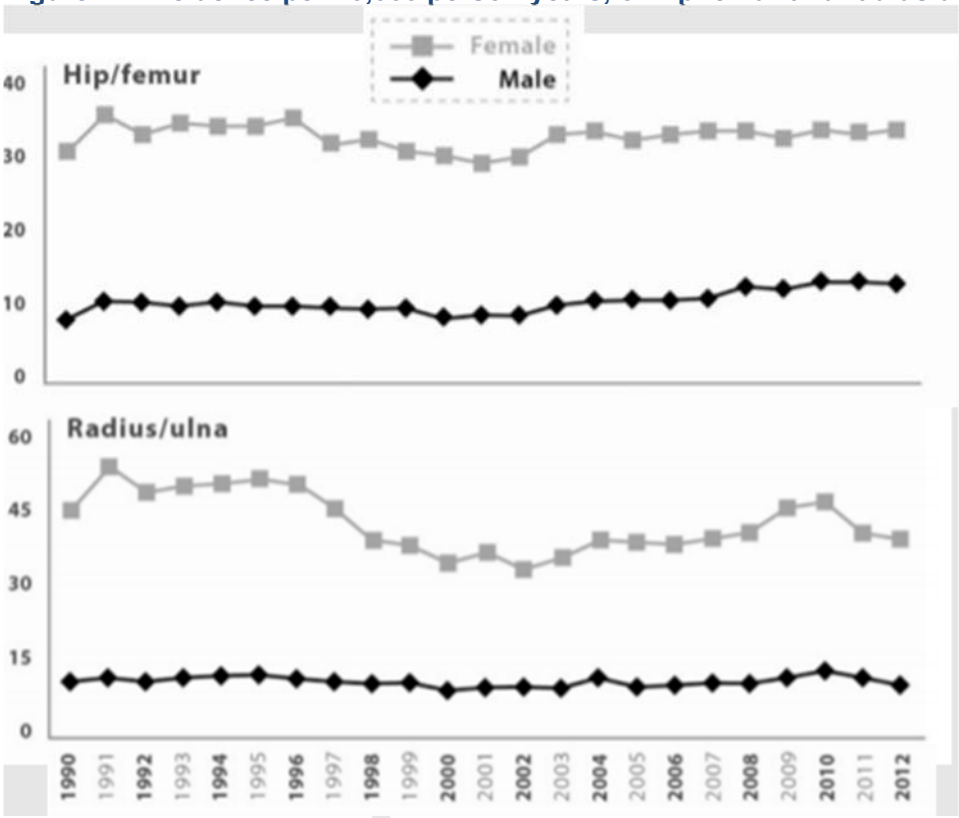
A. On page 78 of the CS it is mentioned that "A retrospective study using the Clinical Practice Research data link (CPRD) in the UK showed that fracture incidences have remained <u>stable</u> over the years 1990–2012 and <u>similar</u> to Singer et al.'s estimates". Please explain (numerically) to what extent fracture incidences have remained stable and similar to those in Singer et al. study.

The incidence of hip/femur and radius/ulna fractures (women and men aged 50 and older), respectively, from 1990 to 2012 are shown in Figure 14 below.[55] The figure shows that the incidence of hip/femur fractures have remained rather stable at about 35 fractures/10,000 person-years over the studied years. Incidence of radius/ulna fractures were slightly less stable over the studied years, with a drop from about 50 fractures/10,000 person-years in 1992–1995 to approximately 40/10,000 person-years in 1998. However, from 1998, the year when Singer et al.'s study was published, to 2012, the in incidence of radius/ulna remained largely unchanged at about 40/10,000 person-years.

The wrist fracture incidence in Singer et al. study in women aged 75–79 was approximately 70/10,000 person-years.[56] In the below study by van der Velde et al., the incidence in the same age group was around 50–70/10,000 person-years, depending on year.[55] The hip fracture incidence per 10,000 in the age group 75–79 was approximately 75 in the Singer et al. study and between 55 and 60 in the van der Velde study.[55, 56] The slightly lower number in the van der

Velde study is expected since the study measured fracture counts retrospectively based on diagnosis coding while the Singer et al study measured fractures based on admission registers and clinical records which may identify more fractures.[55, 56]

Figure 14: Incidence per 10,000 person-years, of hip/femur and radius/ulna in the UK



Source: van der Velde et al. 2016.[55]

B. On page 78 of the CS it is mentioned that "Comprehensive data on the risk of clinical vertebral fractures is limited for the UK, therefore, the UK clinical vertebral fracture incidence was calculated by assuming that the ratio of clinical vertebral fracture to hip fracture in a Swedish-based study is similar to that of the UK":

- ***Please explain (numerically) to what extent data on the risk of clinical vertebral fractures is limited for the UK.***

Unfortunately, there are no published UK data available to inform the risk of clinical vertebral fractures.

- ***Please clarify why the study by Singer et al. has not been deemed appropriate to inform vertebral fractures but it was appropriate for hip and NHNV fractures.***

Underreporting of vertebral fractures, in particular, is a common issue in epidemiological studies. Vertebral incidences from Singer et al. were deemed inappropriate because the reported estimates in Singer et al. are unrealistically low (less than a tenth of figures measured in other

Northern European countries).[57] This is likely because not all vertebral fractures were coded. The authors point out that the incidences of vertebral fractures were lower than other studies and discusses that it could have been due to vertebral fractures being treated in other healthcare facilities than those that were included in the study (i.e., community without reference to orthopaedic trauma service). However, this discrepancy was considered to be specific to vertebral fractures, and therefore did not preclude the derivation of hip and non-vertebral fractures estimates from Singer et al.

- ***Please indicate whether the assumption that the ratio of clinical vertebral fracture to hip fracture in a Swedish-based study is similar to that of the UK has been validated by clinical experts.***

The assumption that the ratio of clinical vertebral fracture to hip fracture has previously been reported in the published literature. In the report by Hernlund et al. in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA), Hernlund et al. also assumed that the ratio of clinical vertebral fracture to hip-fracture in the Swedish-based study is similar to that of the UK in order to derive vertebral fracture incidences for the UK.[49]

This assumption has been shown to hold true in a study by Kanis et al.[16] This study compared the pattern of fractures in Sweden and UK (and USA) which indicated that the relationship between hip, distal forearm and proximal humerus fractures in those countries are very similar.

C. Please explore scenario analyses where vertebral fractures are informed by Singer et al. and where the ratio of clinical vertebral fracture to hip fracture is changed in another (plausible) way.

Results from a scenario analysis using the Singer et al. vertebral fractures incidences are presented in Table 45. However, as discussed in sub-question B above, the vertebral fracture incidences estimates are generally not considered to be reliable. Hence, the scenario analysis should not be considered relevant for decision making in this appraisal, as it likely underestimates the risk of clinical vertebral fractures that are known to have a large impact on costs and QoL, and therefore underestimates cost-effectiveness of romosozumab.

Table 45: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment with vertebral fracture incidences from Singer et al. (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO /ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£30,712
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£9,066

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

B7. PRIORITY QUESTION: Please clarify the following questions regarding <u>risk reduction from treatment</u>:

A. On page 79 of the CS it is mentioned that "Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH". Please provide an example showing how the "continuous hazards approach" was applied.

The continuous hazards approach entailed that hip and non-vertebral fracture data were reconstructed from Kaplan-Meier curves (ARCH CSR, Figure 14-4.3 [page 900] and Figure 14-2.7 [page 907]), for both treatment arms separately. Different parametric distribution functions were then fitted on both datasets (one with romosozumab/alendronate data and one with alendronate) separately, to find the best fit based on Akaike information criterion. Time-dependent rates with the associated hazard function were then calculated separately for both arms (using the mid-point of model cycle). Following this, the hazard ratios of romosozumab vs. alendronate were calculated.

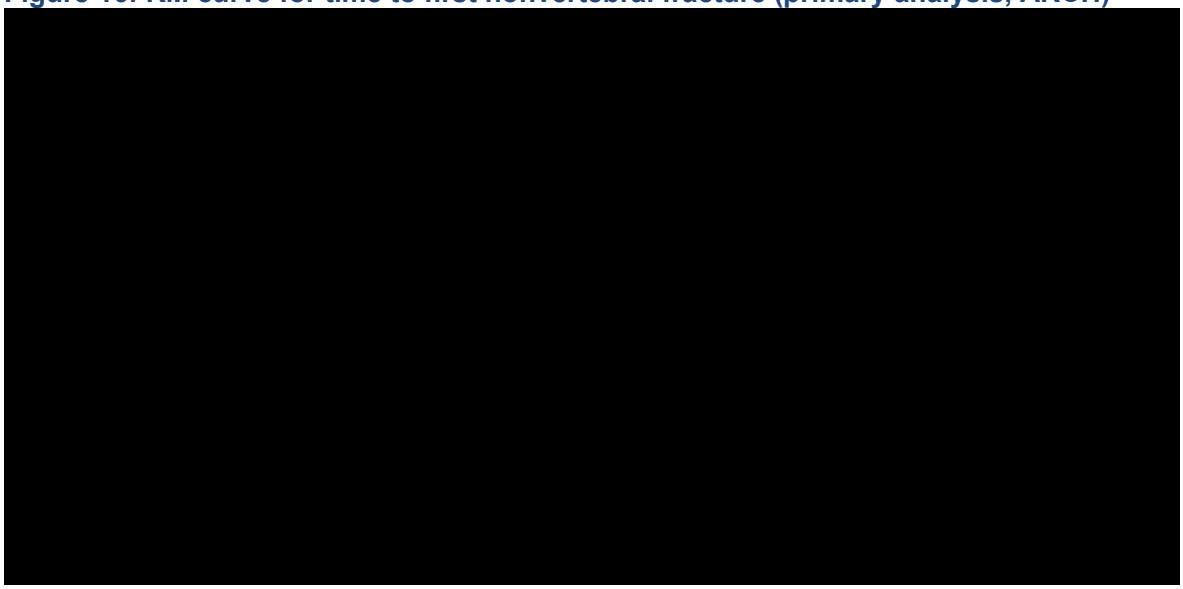
B. On page 79 of the CS it is mentioned that "Patient level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-dependent hazard rates were calculated for the mid-point of the model cycle". Please indicate where these analyses can be found.

These analyses have been conducted internally by UCB and Amgen and are not publicly available. The methods are described in detail in the PowerPoint presentation entitled "B7B PLD KM analyses" as well as in question B7A and B7C. Hip and non-vertebral fracture data were derived from Kaplan-Meier curves from the ARCH study, and parametric functions were fitted to these data as described in detail in question B7C. The Kaplan-Meier curves from the ARCH CSR are presented in Figure 15 and Figure 16.

Figure 15: KM curve for time to first hip fracture (primary analysis, ARCH)

Abbreviations: ALN: alendronate; ALN/ALN: alendronate-to-alendronate; ROMO/ALN: romosozumab-to-alendronate.
Source: ARCH CSR.[32]

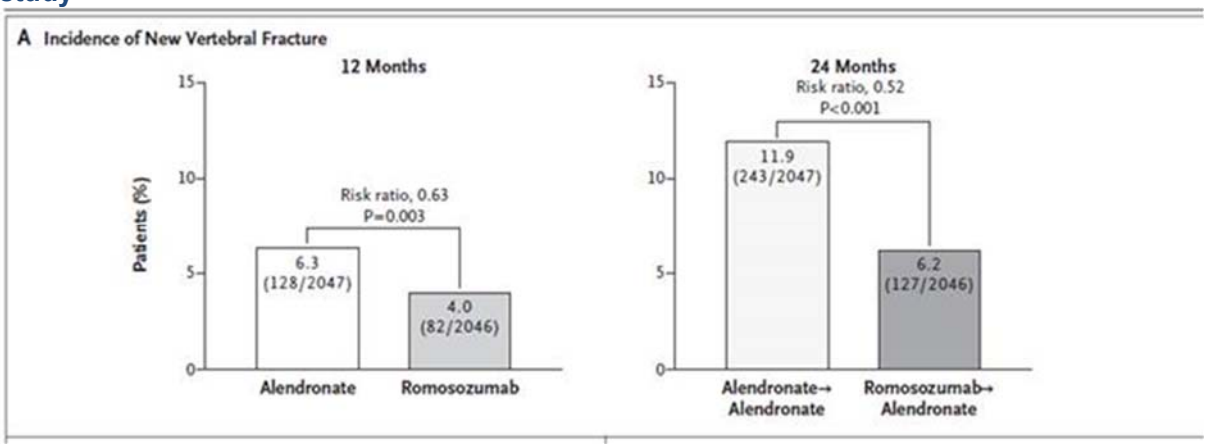
Figure 16: KM curve for time to first nonvertebral fracture (primary analysis, ARCH)



Abbreviations: ALN: alendronate; ALN/ALN: alendronate-to-alendronate; ROMO/ALN: romosozumab-to-alendronate.
Source: ARCH CSR.[32]

New vertebral fracture data were calculated from published data (Figure 17).[30] An example of how the non-cumulative efficacy for vertebral fractures is described in Table 46 below.

Figure 17: Incidence of new vertebral fractures at Month 12 and Month 24 in the Saag et al. study



Source: Saag et al. (2017).[30]

Table 46: Calculation of non-cumulative efficacy for vertebral fractures

Time/Nr of fx	Alen+alen N=2047	Romo+alen N=2046	RR
0	0	0	-
12 mo	128	82	=(82/2046) / (128/2047)
24 mo	243	127	=((127-82)/(2046-82))/((243-128)/(2047-128))

Abbreviations: Alen: alendronate; Fx: fracture; Romo: romosozoumab; RR: relative risk

C. Please explain how the hazard ratios shown in Table 19 were calculated. Please indicate what hazard ratios are used in the model after 36 months (until the end of the time horizon).

The hazard ratios in Table 19 were calculated based on the methods described on page 79 of the CS Document B. Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH. Patient level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-dependent hazard rates were calculated for the mid-point of the model cycle. In the model, efficacy of non-vertebral fractures was applied to NHNV fractures due to lack of data on all fractures excluding both hip and vertebral. For vertebral fractures, efficacy of new vertebral fractures was calculated from the published data at 12 and 24 months. Efficacy for vertebral fractures beyond month 24 is based on 24 month- efficacy. Hazard ratios for treatment beyond 36 months (i.e., month 37 to month 60) were based on the 36-month efficacy.

D. Please justify (both numerically and conceptually) why HRs (from Table 19) and RRs (from the NMA) "give practically the same information".

Although some technical differences exist, hazard ratios and relative risks are conceptually similar as they are both a relative measure of disease occurrence.[58] They primarily differ in terms of time period, while relative risks are cumulative over time (typically over the study period),

hazard ratios represent the instantaneous risk over the study period (or the difference in risk at any particular time during the study period).

As an example, both the hazard ratio and the relative risk of hip fracture for romosozumab/alendronate vs alendronate were 0.62 at the primary analysis in ARCH. Furthermore, the relative risk on non-vertebral fracture was 0.82 and the hazard ratio 0.81.[30] Given these similarities and given the lack of time-dependent relative risks from ARCH, it was deemed reasonable to use relative risk and hazard ratios interchangeably.

E. On page 79 of the CS it is mentioned that "The approach of using the alendronate vs. placebo data is reasonable given that the efficacy data of alendronate vs placebo from UCB's NMA do not differ significantly from other NMAs, for example NICE's most recent NMA (Table 20)". We consider this statement rather subjective seeing the values presented in Table 20. This is particularly the case for the values shown for teriparatide, which is the most effective treatment according to the AG NMA but not in the company's NMA. This raises concerns about the validity/credibility of the NMA results. Please provide separate results based on either NMA.

The results for alendronate vs placebo were similar in both the NICE NMA and the UCB NMA. This provides further validation for using the efficacy of alendronate vs placebo to calculate the efficacy of romosozumab vs placebo, in the absence of relevant data from the ARCH trial. A scenario analysis where the efficacy of alendronate vs placebo for hip, vertebral and NHNV fractures are based on NICE's NMA demonstrates similar results to UCB's NMA (Table 47).

UCB acknowledge that teriparatide had a better effect vs placebo in NICE's NMA compared with UCB's NMA for hip and other fractures, while it was similar for vertebral fractures. However, it is important to consider that while most of the RCTs included in the NICE NMA were conducted in postmenopausal women, there were some trials of men and some including patients with steroid induced osteoporosis.[33, 34] This means that the evidence bases for the NICE and UCB NMA's are different. As such, cost-effectiveness scenarios utilising the NICE NMA is not appropriate for decision making, as the underlying evidence base is outside the licensed indication for romosozumab.

Table 47: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment using efficacy of alendronate vs placebo from NICE's NMA (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO/ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,902
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£4,219

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

B8. Please provide new versions of Table 21 and 22 showing risk ratios for the complete modelled time horizon.

Table 21 and 22 showing risk ratios up until the maximum treatment duration (60 months) are included below. No treatment efficacy is modelled (i.e., the risk ratio =1) after the treatment + offset period and therefore not shown in the tables.

Table 48. Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and NMA (ITT populations)

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Abbreviations: CI: confidence interval; ITT: intention-to-treat; NMA: network meta-analysis; NHNV: non-hip, non-vertebral.

Table 49. Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and scenario NMA (EU label-matched population)

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Abbreviations: CI: confidence interval; ITT: intention-to-treat; NMA: network meta-analysis; NHNV: non-hip, non-vertebral.

B9. PRIORITY QUESTION: Please clarify the following questions regarding <u>modelling of persistence</u>:

A. Please indicate the main causes for treatment discontinuation as observed in the ARCH trial and in the UK study by Li et al. 2010.

Table 50 shows the reasons for discontinuation of the full analysis set at the time of the primary analysis of the ARCH trial.

Table 50: Patient disposition (full analysis set, primary analysis)

	Alendronate 70 mg QW/ Alendronate 70 mg QW (N = [redacted]) n (%)	Romosozumab 210 mg QM/ Alendronate 70 mg QW (N = [redacted]) n (%)	All (N = [redacted]) n (%)
Double-blind period accounting			
Completed double-blind period	[redacted]	[redacted]	[redacted]
Discontinued study during double-blind period	[redacted]	[redacted]	[redacted]
Consent withdrawn	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Noncompliance	[redacted]	[redacted]	[redacted]
Ineligibility determined	[redacted]	[redacted]	[redacted]
Protocol deviation	[redacted]	[redacted]	[redacted]
Administrative decision	[redacted]	[redacted]	[redacted]
Requirement for alternative therapy	[redacted]	[redacted]	[redacted]

Abbreviations: QM: once monthly; QW: once weekly.

Unfortunately, the reasons for patients discontinuing from treatment in the UK study by Li et al. (2012) are neither presented in the Li et al. (2012) publication nor held by UCB and therefore cannot be presented.

B. On page 84 of the CS it is mentioned that "Treatment discontinuation resulted in patients not receiving the same anti-fracture benefits as would be expected for a fully persistent patient (i.e., a patient still on treatment)". Please clarify whether the "effects" of treatment discontinuation have also been included in the costs and quality of life sides of the economic analyses.

The effects of treatment discontinuation on costs and QoL are included by the increase of fracture risk when stopping treatment early compared with completing the treatment. After stopping treatment, anti-fracture efficacy is lost (by decreasing over a time period equal to the

treatment length due to residual effect). No other effects of discontinuation are included in the model since there is no evidence that discontinuation per se has an impact on costs or effects. Also, such effects have to our knowledge not been included in prior economic evaluations. For example, it was not included in the recent economic evaluation of non-bisphosphonates by NICE.[33]

C. Please justify the assumption that patients are at risk of dropping out <u>during the first three years</u>. Please clarify whether this assumption is applied to all treatments, regardless of the sequence. For example, for the intervention romosozumab (ROMO) + alendronate (ALN), patients are at risk of dropping out during the first three years in total (1 year of ROMO and 2 years of ALN) or during the first three years per treatment (1 year of ROMO and 3 years of ALN – so 4 years in total). In any case, this does not seem to match with the values shown in Table 25 where discontinuation is possible for some treatments up to year 5.

UCB have double-checked the values in Table 25 of the CS and can confirm that they are correct. However, following discussion with the ERG and NICE, UCB understand that the description of persistence in the company submission is unclear. UCB can confirm that in the cost-effectiveness model, patients are at risk of dropping out during the entirety of the treatment duration, and not solely during the first three years.

D. Please clarify what happens to patients after dropping out of one treatment: do they switch to the next in the sequence or do they all go to placebo? Please justify this assumption.

Patients who drop out of treatment do not switch treatment but remain without treatment for the remaining time horizon. This assumption was made because there are no data available on switching patterns for sequential treatments and it is likely to have limited impact on the results. Regardless the same assumptions were applied to all treatments,

E. On page 84 of the CS it is mentioned that "In the base case cost-effectiveness analysis, persistence on alendronate alone (when not preceded by romosozumab) was derived from Li et al. (2012), a UK General Practice Research Database (GPRD) of persistence on osteoporosis medications among postmenopausal women in the UK". Please clarify why this was not based on ARCH data. Please provide a comparison between persistence estimates in Li et al. and the ARCH trial.

Persistence data from ARCH was not used in the model, because persistence data from retrospective observational studies are more appropriate than persistence data from clinical trials. Persistence in clinical trials is significantly higher than in clinical practice most likely

because patients are being observed and have consented to participate in the study. The guidelines from 2019 for the conduct of economic evaluations in osteoporosis endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the IOF recommends using real-world data on medication adherence.[59]

Table 51: Proportion of patients on alendronate based on Li et al. and ARCH trial

Month since treatment initiation	Alendronate alone based on Li et al. (2012)[60]	Alendronate alone based on ARCH.[30]
6	49%	Not available
12	38%	[redacted]
18	34%	Not available
24	30%	[redacted]
30	27%	Not available
36	24%	Not available
42	22%	Not available
48	20%	Not available
54	19%	Not available
60	17%	Not available

Source: Li, L., et al. (2012).[60]; Supplement to: Saag et al. (2017).[30]

F. There are several sources of uncertainty regarding persistence on romosozumab in UK clinical practice and the company has acknowledged that this is still unknown. However, there are certain assumptions that require further justification:

i. As a starting point, a Swedish study reporting persistence on teriparatide has been used. Please indicate whether it was not possible to use UK studies for this. In case it was not, please justify that the Swedish study is representative for the UK.

There is a paucity of UK persistence data for teriparatide, and therefore for the reasons cited in previous questions regarding the generalisability of Swedish data to the UK, the use of the Swedish study was considered to represent the best proxy for persistence data for teriparatide in UK clinical practice.

ii. The company stated that since romosozumab will be administered much less frequently compared to teriparatide, it is reasonable to assume that patients treated with romosozumab will exhibit higher persistence compared with teriparatide. While this might be the case, it might also be possible that patients could discontinue romosozumab for other reasons. Please justify this assumption.

Romosozumab is administered much less frequently compared to teriparatide: romosozumab is administered once every month, for 12 months as two subcutaneous injections, resulting in a total number of 24 injections over one year. In comparison, teriparatide is administered once daily for two years, consisting of approximately 730 injections over the two-year period.

Insights collected from UK HCPs have highlighted the importance of patient choice when selecting a suitable therapy. Fewer injections and the availability of a patient support programme (PSP) were important considerations when considering therapy and gauging potential persistence on treatment.

In Northern Ireland, UCB were informed that teriparatide does not offer a comprehensive PSP to include the demonstration/support with the injections. This greatly impacts the patient's confidence and adherence in taking the drug. Additional insights were previously discussed as part of the NICE scope consultation.

iii. **The magnitude of the improvement in persistence on romosozumab is unknown. The estimated persistence was estimated from clinical trial data. It might be expected that persistence is higher in clinical trials than in daily practice. That might be the reason why persistence on alendronate alone was derived from Li et al. (2012) and persistence on teriparatide was derived from the Swedish osteoporosis database. If that's the case, this approach (using trial data for romosozumab only) would be inconsistent and most likely biased in favour of romosozumab. Also, the assumptions made on page 85 of the CS "For the treatment sequence of romosozumab followed by alendronate used in this submission, [redacted]. This is based on the assumption that patients who have initially demonstrated high persistence on romosozumab would be expected to demonstrate high persistence on follow-on treatments, and therefore the persistence on alendronate after romosozumab would be [redacted] than the persistence on alendronate alone reported by Li et al. (2012)"; are not justified enough. For those reasons, the estimates provided on Table 25 are uncertain, some of them inconsistent/unjustified and likely to favour romosozumab. Therefore, in any case, [redacted], where:**

- **persistence estimates for all treatments are based on trial data (even though this would most likely overestimate persistence for all treatments);**

- **persistence estimates for romosozumab are equal to persistence estimates for teriparatide (even though it might be expected that for romosozumab these would be higher – this could be seen as a conservative approach); and**

- **persistence is 100% (no treatment discontinuation).**

Persistence data from retrospective observational studies are more appropriate than persistence data from clinical trials. Persistence in clinical trials is significantly higher than in clinical practice most likely because patients know they are being observed and have consented to participate in the study. The guidelines from 2019 for the conduct of economic evaluations in osteoporosis endorsed by the ESCEO/IOF recommends using real-world data on medication adherence.[59]

Persistence of romosozumab is assumed to be the same as in the ARCH trial, despite clinical trials show higher persistence than what is seen in clinical practice. This was necessary given that there is no real-world evidence currently available for romosozumab as it has only been recently launched. This assumption around persistence is reasonable given that UCB will introduce a new Patient Support Programme (PSP). The PSP will provide a homecare service and offer the option for patients to join an adherence support programme. This will ensure that patients adhere to romosozumab, and then successfully transition to a follow-on therapy after one year of treatment with romosozumab. The PSP will be provided alongside and extend beyond the romosozumab treatment period for up to 15 months to ensure the transition to the recommended follow-up treatment. With the support from the PSP, it is reasonable to believe that patients will be able to follow the treatment regimen to a similar extent as in the trial.

A scenario has however been tested where persistence to romosozumab/alendronate and alendronate alone are based on the ARCH trial. Persistence at 12 months for romosozumab was [redacted] ([redacted] of [redacted] patients completed the 12-month double-blind period) and [redacted]% ([redacted] of [redacted] patients) for alendronate. Persistence at 24 months was [redacted]% ([redacted] out of [redacted] completed the primary analysis period) for romosozumab/alendronate and [redacted]% for alendronate alone ([redacted] out of [redacted] patients). Persistence beyond the primary analysis period (month 30 to 60) was linearly extrapolated based on the drop-off rate between month 12 and 24 and 6/18 months (Table 52).

Table 52: Persistence based on ARCH (linearly extrapolated beyond trial follow-up)

Persistent proportion	Romosozumab/alendronate	Alendronate
6	[redacted]	[redacted]
12	[redacted]	[redacted]
18	[redacted]	[redacted]
24	[redacted]	[redacted]
30	[redacted]	[redacted]
36	[redacted]	[redacted]
42	[redacted]	[redacted]
48	[redacted]	[redacted]
54	[redacted]	[redacted]
60	[redacted]	[redacted]

The results presented in Table 53 are not considered relevant for decision making as persistence inputs derived from clinical trial settings are known to differ substantially from real world persistence of osteoporosis patients and are at high risk to misrepresent the cost-effectiveness of romosozumab.

Table 53: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment with persistence data based on ARCH for all treatments (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO /ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£54,340
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£646

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

Persistence to romosozumab is unlikely to be equal to teriparatide's persistence given that romosozumab is given in monthly intervals and teriparatide is given in daily intervals (only 24 romosozumab injections are needed while teriparatide requires 720 injections over the treatment course). Longer durations between administrations are known to be associated with better persistence compared with shorter time between administrations.[61] Whilst romosozumab will be offered with a PSP to ensure persistent use of romosozumab and transition to follow-on therapy, such a support program is not available for the support with the injections of teriparatide to UCB's knowledge.

Despite the important aforementioned concerns relating to the questionable validity and relevance for decision-making of such a scenario assuming equal persistence of romosozumab to teriparatide, the results from this scenario are presented in the table below. Persistence for teriparatide is based on the same source as described in the CS.

Table 54: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment with romosozumab's persistence equal to teriparatide's persistence (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO/ ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£38,295
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£10,016

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

Issues with persistence to osteoporosis treatments are well-known, and persistence to osteoporosis treatments is highly unlikely to be 100% for any drug. Such scenario is therefore unrealistic but the results assuming no discontinuation for romosozumab/alendronate and

alendronate alone are shown below and demonstrate romosozumab/alendronate to be cost-effective against alendronate alone.

Table 55: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment with 100% persistence for all treatments (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO /ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£20,989
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Cost-saving

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

However, it is important to reiterate that the scenarios presented in response to this question should be interpreted with caution, considering the substantial limitations with regards to the validity and relevance for decision making associated with the modelled assumptions and resulting cost-effectiveness estimates. Nevertheless, these scenarios demonstrate that romosozumab remains a cost-effectiveness treatment option versus no treatment in all of the additional scenarios, as well as being cost-effective versus alendronate in one of the three extreme scenarios considered.

B10. PRIORITY QUESTION: Please clarify the following questions regarding dynamic residual effects:

A. Please provide numerical examples illustrating how dynamic residual effects are included in the model.

The tables below show two examples of how dynamic residual effects are calculated. Example A shows how treatment effect for a patient who is fully persistent and thereby completes the 12-month romosozumab treatment and the subsequent 48 months with alendronate. This patient has full effect (i.e., the relative risk of romosozumab vs placebo based on the trials and NMA described on page 81 in the CS) of treatment until the start of year 6. From that point, treatment effect decreases linearly to no effect when additional 5 years have passed. This patient has therefore full effect in 5 years, and 5 years of decreasing effect from treatment. The treatment effect during the residual period, is calculated based on the treatment effect in the last cycle of treatment and adjusted for an effect multiplier that linearly decreases to 0 until the residual effect period ends.

Example B show an example for a patient who stops treatment early, at end of year 1. This patient has full treatment effect in year 1. From the start of year 2, treatment effect linearly decreases to no effect when 1 year has passed. This patient has full effect for 1 year and residual effect for 1 year. Treatment effect during the residual period is calculated in the same way as in example A (Table 56); treatment effect is based on the effect in the last treatment cycle (cycle 2) and adjusted for the effect multiplier.

Table 56: Example A: Calculation of dynamic residual effects for a patient who is fully persistent (completes the 12-month romosozumab period and 48 months of alendronate)

Cycle (length 6 months)	Effect multiplier	Hip fracture treatment effect (RR romosozumab vs. placebo)	Comment
1	1	[redacted]	Patient on treatment. Full treatment effect (effect multiplier=1).
2	1	[redacted]	
3	1	[redacted]	
4	1	[redacted]	
5	1	[redacted]	
6	1	[redacted]	
7	1	[redacted]	
8	1	[redacted]	
9	1	[redacted]	
10	1	[redacted]	
11	0.95	[redacted]	Patient has stopped treatment. Effect multiplier linearly decreases to 0 until 5 years (=treatment length) has passed. The treatment effect is based on the effect in the last cycle of treatment, here [redacted] in cycle 10, and adjusted for the effect multiplier
12	0.85	[redacted]	
13	0.75	[redacted]	
14	0.65	[redacted]	
15	0.55	[redacted]	
16	0.45	[redacted]	
17	0.35	[redacted]	
18	0.25	[redacted]	
19	0.15	[redacted]	
20	0.05	[redacted]	

Abbreviations: RR: relative risk.

Table 57: Example B: Calculation of dynamic residual effects for a patient who stops treatment with romosozumab early, at end of year 1

Cycle (length 6 months)	Effect multiplier	Hip fracture treatment effect (RR romosozumab vs. placebo)	Comment
1	1	[redacted]	Patient on treatment. Full treatment effect (multiplier=1)
2	1	[redacted]	
3	0.75	[redacted]	Patient has stopped treatment. Effect multiplier linearly decreases to 0 until 1 year (=treatment length) has passed. The treatment effect is based on the effect in the last cycle of treatment, here [redacted] in cycle 2, and adjusted for the effect multiplier
4	0.25	[redacted]	
5	0	[redacted]	2 years have passed (1 year treatment + 1 year offset time), no treatment effect for the rest of the time horizon
6	0	[redacted]	
7	0	[redacted]	
8	0	[redacted]	
9	0	[redacted]	

10	0	[redacted]	
11	0	[redacted]	
12	0	[redacted]	
13	0	[redacted]	
14	0	[redacted]	
15	0	[redacted]	
16	0	[redacted]	
17	0	[redacted]	
18	0	[redacted]	
19	0	[redacted]	
20	0	[redacted]	

Abbreviations: RR: relative risk.

B. Please define also what is meant by "partially persistent patients" and include these patients in the numerical examples.

We acknowledge the term "partially persistent" may be unclear. Partially persistent refers to patients who drop out of treatment before the intended treatment length ("non-persistent"). An example of how residual effect is calculated is given in Example B in sub-question A above. Dynamic residual effect is calculated in the same way for fully persistent and patients who stops treatment before the intended treatment length, but the period of residual effect is adjusted to the actual time on treatment for the individual patient as explained in greater detail in response to question B10-A.

B11. PRIORITY QUESTION: Please clarify the following questions regarding <u>modelling of mortality</u>:

A. Please provide a numerical example showing how the three mortality rates mentioned in the CS (age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor) are included in the model.

Mortality rates for a patient after a disease event consist of 1) the age specific mortality of the general population for the general population (all-cause mortality); 2) a relative risk capturing the excess mortality of the disease; and 3) a co-morbidity adjustment factor. The co-morbidity adjustment factor considers the possibility that excess mortality among the patients with a specific disease is not entirely attributable to that disease. This approach is recommended by ESCEO/IOFs recommendations for the conduct of economic evaluations in osteoporosis.[59]

The mortality rate following a hip fracture in a patient who is 80 years old at the time of the hip fracture is calculated as follows. The age specific mortality of the general population for the general population (all-cause mortality) for an 80-year-old woman in the UK is approximately 0.04.[62] The increase in risk of death in the year after a hip fracture is 2.92 in an 80-year-old woman. The co-morbidity adjustment factor is assumed to be 30% in line with ESCEO/IOF recommendations,[59] previous health economic studies[63, 64], and studies by Parker and Anand and Kanis et al.[65, 66]

The risk of death in this hypothetical patient is then calculated by multiplying the normal age and gender specific mortality with the relative risk of death in a patient with hip fracture adjusted for the co-morbidity adjustment factor: 0.04*2.92*(1-30%)=0.08.

B. On page 87 of the CS it is mentioned that "All patients are at risk of dying corresponding to the risk of the UK general population from the start of the model". Please clarify why at the start of the model the risk of dying is not that of the patient population.

The risk of death is based on the UK general population mortality, but is subsequently adjusted for the risk factors of the patient population under consideration. This means that patients who have a recent fracture at treatment start (i.e., model start) are also assumed to have increased risk of dying according to the relative risk capturing excess mortality of fracture, adjusted for the co-morbidity adjustment factor (30%).

Furthermore, risk of death is adjusted for other clinical risk factors (CRFs) based on FRAX. Some of the FRAX CRFs contribute to mortality and FRAX outputs the relative risk of pre-fracture mortality dependent on the defined patient population. This relative risk should be used to adjust the baseline mortality in the cost-effectiveness analysis and the mortality after fracture. However, it should be acknowledged that this assumes that the relative risk of mortality that is obtained from FRAX which is related to patients that will and will not fracture is maintained after fracture. This assumption is made since the relationship between the CRFs and the risk of mortality after fracture is not yet investigated. The main consequence of using the FRAX mortality relative risks is that high risk populations will have a higher overall mortality and thus benefit less from avoiding fractures, compared to modelling without the mortality adjustment. Higher overall mortality leads to higher incremental life years gained from treatment. However, higher overall mortality has a negative impact on QALYs gained because quality-of-life impact of a fracture is relative to not having a fracture and with lower expected life years in both treatment arms, the impact of the fracture on quality-of-life decreases.

FRAX adjusts only mortality related to the CRFs and other factors that might differentiate the mortality in osteoporosis patients compared to the general population are not accounted for. Therefore, the assumption that only a proportion of the excess mortality after fracture can be related to the fracture event is retained. The model uses the highest mortality in situations where both post-fracture mortality and FRAX-derived mortality should be accounted for.

C. Please justify the choice of 30% relative risk of death associated to a fracture compared to no fracture (CS pages 87-88).

Patients with osteoporosis have a higher degree of frailty compared to the general population and excess mortality after a fragility fracture is not entirely attributable to the fracture event. A common assumption[63, 64]in health economic studies is that 30% of the excess mortality is directly caused by the fracture and this is supported by studies by Parker and Anand and Kanis et al.[65, 66] However, other studies[21, 67, 68] claim that there is little or no relation between co-morbid conditions and post-fracture mortality, which consequently would imply that more than 30% of the excess mortality is caused by the fracture itself. The ESCEO/IOF recommendations of economic evaluations in osteoporosis recommends that mortality should be adjusted by 25-30%.[59] Thus, in agreement with previous health economic studies, it was assumed that 30% of excess mortality after a hip, vertebral and non-hip-non-vertebral fracture is associated with the fracture event.

D. Please justify the assumption that "the standardised mortality ratios (SMRs) estimated using the Swedish data would be generalisable to the UK due to the similarity in access to health care between the two countries" (CS page 88). Please conduct scenario analyses where this SMR is varied within a plausible range of values.

As noted in Question B4, D, due to a lack of comparable studies in other countries that have estimated the relative risk of subsequent fractures, it is difficult to conclude with certainty that the relative risks would be exactly the same in Sweden and the UK. However, for the same reasons discussed in Question B4, D, it is reasonable to consider that SMRs estimated using Swedish data would be generalisable to the UK.

Similar assumptions have been made in the published literature. Hernlund et al, in a study endorsed by the International Osteoporosis Foundation, applied Swedish relative risks of death to compute absolute risk of death after fracture in other countries, such as the UK.[49] Previous cost-effectiveness studies have made similar assumptions on fracture risk and mortality after fracture when there was a lack of country-specific data.[50]

A scenario analysis has been conducted where the relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al (UK setting).[69] The excess mortality rates from van Staa (Table 58) were transformed to relative risks by applying the rates to the general population mortality and dividing that by the general population mortality. The relative risks in the second and following years for hip and vertebral fractures, and first year for non-hip-non-vertebral fractures, were assumed to be the same as in the base case (as described on page 88 of the CS). This change had a minor impact on the cost-effectiveness results.

Table 58: Excess mortality rates from the van Staa study for hip and vertebral fractures during the first year after fracture used in the scenario analysis

Age (year)	Hip (%)	Vertebral (%)
50–59	2.4	2.3
60–69	4.4	3.5
70–79	7.5	5.2
80–89	11.4	6.7
90+	13.6	6.6

Table 59: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment using relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al. (2007)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO/ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£16,728

No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£3,801

Source: van Staa et al. (2007).[69]
Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

E. On page 88 of the CS, it is mentioned that "As the variation in fracture distribution was not considered to be large across different age groups, the same relative risk was used for all ages". Please provide evidence to support this assumption. Also, please explain why "Using the same relative risk after NHNV fractures for all ages could thus possibly underestimate mortality in younger patients and overestimate mortality in older patients".

The relative risks of death after NHNV fractures were based on a study by Barrett et al.[70] The specific relative risks for NHNV were not presented by age in the paper, however, the authors note in the discussion that there was no significant change in the relative risk depending on age for NHNV fractures. Therefore, it was assumed that the relative risk of death was the same regardless of age (weighted relative risk of 1.23, see Table 27 in the CS).

Using the same relative risk after NHNV for all ages, i.e., the mid-point for all ages, could potentially underestimate the mortality in younger patients and overestimate mortality in older patients because the increase in risk after a fracture is generally higher in younger patients. For example, the relative risk after hip and vertebral fracture is highest in the youngest women and decreases with higher age; the relative risk after hip fracture is 9.8 in a 50-year-old woman and 1.63 in a 90-year-old woman.

Adverse events

B12. PRIORITY QUESTION: P90 of the CS states "an imbalance in serious adjudicated cardiovascular (CV) adverse events (AEs) was observed in the ARCH trial. As a result, romosozumab is contraindicated for patients with previous myocardial infarction or stroke. Given this contraindication, which was not an exclusion criterion in the ARCH trial, it was considered reasonable to exclude CV AEs from the economic analysis".

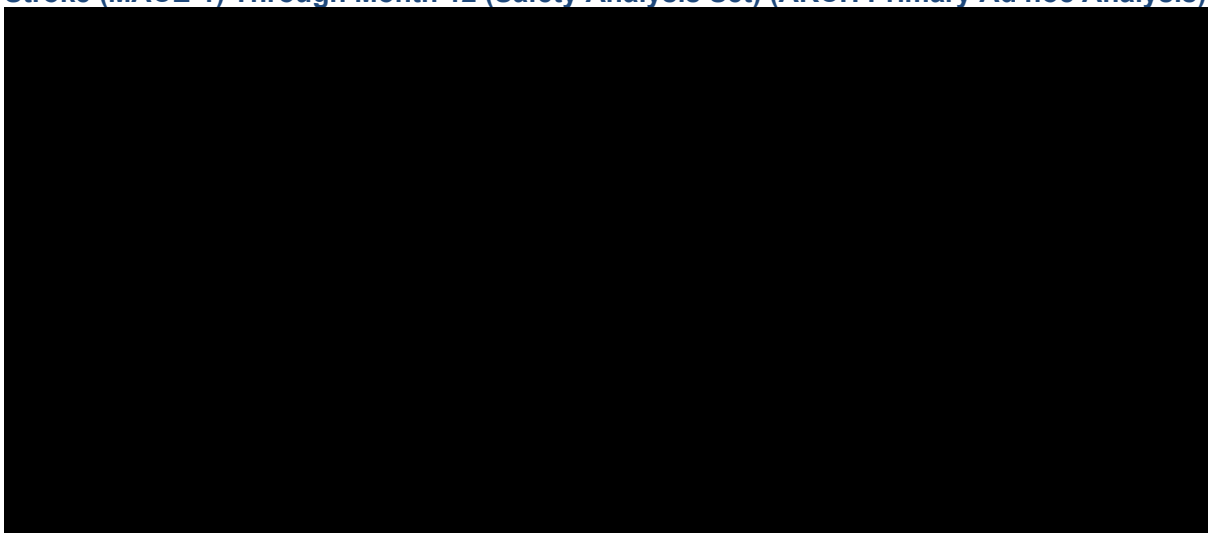
- **Please conduct an analysis showing the proportion of people who experienced a CV AE in the ARCH trial who had a history of myocardial infarction or stroke.**

- **Please include an option in the model to include CV AE according to the incidence in the ARCH trial and relevant disutilities and costs.**

A subgroup analysis of the treatment-emergent adverse events (TEAE) for major cardiovascular events (MACE) is presented in Figure 18 below from the romosozumab EPAR.[71] This post hoc

safety analysis showed a trend toward a higher risk of MACE in patients with a history of CV events.[71]

Figure 18: Medical History Subgroup Analyses: Time to First Occurrence of Positively-adjudicated Cardiovascular Event Leading to Death, Serious Myocardial Infarction or Stroke (MACE-1) Through Month 12 (Safety Analysis Set) (ARCH Primary Ad hoc Analysis)



It is important to note that the strongest predictor of a subsequent MI or stroke is a recent MI or stroke. The risk of a further event is highest in the first year and continues to decrease with time, thereafter. The proposal to contraindicate romosozumab in all patients with a history of MI or stroke regardless of when those events occurred represents a conservative approach to manage the absolute risk of CV events.

The contraindication to exclude patients with a history of MI and stroke resulted from a conservative position, supported by the EMA, despite the fact that no causality could be established between romosozumab and the observed imbalance in CV events. Extensive data analyses as well as pre-clinical studies could not identify single or combined risk factors to identify patients at increased relative risk of CV AEs. Thus, in order to minimise the risk at the population level the regulatory decision was made to contraindicate romosozumab in the group with the highest incidence of MACE events regardless of treatment, i.e. those with a history of MI and stroke. Because of the lack of identified plausibility for the recorded imbalance observed in ARCH, a post-approval safety study is in place to further characterize the use of romosozumab and CV events on an ongoing basis.

The base case cost-effectiveness in the CS did not account for CV cost- and health effects due to FRAME and other romosozumab studies not revealing any imbalances. Table 60 below presents results from the requested scenario analysis where the impact of CV events on costs and QoL has been considered (this option has also been included in the cost-effectiveness model). This scenario using ARCH CV rates represents a very conservative approach as no imbalance in serious CV events was noted in the larger placebo-controlled study (FRAME). The relative risk of a CV-event was based on the ARCH study, including only patients who do not have the contraindication of prior myocardial infarction or stroke.

Post-hoc analyses of ARCH showed that patients randomized to romosozumab and did not have the contraindication (MI or stroke) at baseline, had a relative risk of major adverse cardiovascular events (MACE) of [redacted] during the first [redacted] years after randomization, compared with alendronate (subject incidence [redacted]% in romosozumab arm vs. [redacted]% in alendronate arm).

The risk of a CV-event was based on several sources. Incidences of stroke, MI, angina, coronary insufficiency and venous thromboembolism were pooled from various sources in the published literature.[72-75] A multiplier for QoL after a CV-event was based on a Swedish study by Lindgren et

al.[76] which estimated a quality-of-life loss of 0.075 (multiplier 0.910) during the first year after CV event. For the second and following years, the multiplier was assumed to be 0.95 due to lack of data.

The relative risk of death compared with the general population was calculated by pooling data from the Swedish patient registry of mortality after stroke, VTE, angina, acute heart failure and other CV events (unpublished data). The use of relative risks and multipliers from Swedish sources is supported by that the countries are similar, and the relative impact of CV events is unlikely to differ.

A systematic literature review from 2018 identified one UK study of direct costs related to CV events.[77] This study by Danese et al. estimated hospitalisation costs, outpatient referrals, primary care visits and medications of MI, stroke, unstable angina, heart failure, transient ischemic attack, and coronary artery bypass graft/percutaneous transluminal coronary angioplasty (CABG/PTCA), using HES and CPRD data.[78] The estimated mean costs in month 1-6 after the first CV event was £4594.16 in 2014 prices (£4993.85 in 2020, inflated using the indexes in Table 63). Mean annualised cost in month 7–36 was £2262.92 in 2014 prices (inflated to £2459.79 in 2020 prices). The economic model was built to accommodate first and subsequent year costs, respectively. Therefore, the estimated CV treatment costs by Danese et al., in month 1–6 were applied in the first year and the costs in month 7–36 were applied annually in every subsequent year until end of model time horizon or death (conservative approach). The first-year cost may therefore be slightly overestimated in the model, since the majority costs likely occur closely to the event, which could be considered a conservative approach.[79]

The results of this scenario can rightfully be considered conservative for romosozumab as the CV occurrence rates for romosozumab and alendronate were chosen from the study where the imbalance between these two treatments was greatest (ARCH) and subsequent year costs are applied every year after the CV event until the end of the modelled time horizon or death. The decision not to select or pool any other romosozumab studies (FRAME, STRUCTURE, McClung) where the CV event rate for romosozumab was lower than in ARCH to derive cost-effectiveness results of this scenario means that the results should be considered to be extremely conservative, and for illustrative purposes only. If the CV events for romosozumab were sourced from FRAME or any other study, the ICER would most likely be lower than the ICER presented using ARCH CV event rates.

Table 60: Scenario results for pairwise comparison of romosozumab/alendronate versus alendronate or no treatment including CV events (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£) ROMO/ALN vs comparator	Incremental LYG ROMO/ALN vs comparator	Incremental QALYs ROMO/ALN vs comparator	Pairwise ICER (£/QALY) ROMO/ALN vs comparator
ROMO/ALN	[redacted]	[redacted]	[redacted]				
ALN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£19,500
No treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£5,075

Abbreviations: ALN: alendronate; ICER: incremental cost-effectiveness ratio; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate.

B13. PRIORITY QUESTION: Please justify why only gastrointestinal adverse events (AEs) are included in the model and provide the option in the model to

include all AEs at or above a 5% incidence threshold for either treatment arm for all Grade 3 or higher AEs.

Other AEs were not included in the model since no imbalances, except for CV events, were seen in the ARCH trial. ARCH reported that serious AEs occurred in 12.8% for romosozumab vs. 13.8% for alendronate during the 12-month double-blind period.[30]

At 24 months, the proportions with serious AEs were 28.7% for romosozumab vs. 30.0%. The estimated between-group differences were not statistically significant (p>0.05). Gastrointestinal AEs were however included since this is a common AE of oral bisphosphonates and is often included in models of bisphosphonates. For example, the independent academic Assessment Group included effect on QoL and costs due to GIAEs in the recently suspended assessment of non-bisphosphonates.[33] The independent assessment group modelled AEs of osteonecrosis of the jaw and atypical femoral fractures, however, the incidence in the ARCH trial and difference between the treatment arms were very small (<0.2% for romosozumab/alendronate and alendronate alone). Given the small differences between the treatment arms, and in order to align with the approach of the independent academic Assessment Group, the option to consider additional AEs has not been included in the model.

Health-related quality of life

B14. PRIORITY QUESTION: Were the utility multipliers from the ICUROS study based on data from all countries in the dataset, a subset of countries or UK-specific? Please also justify your choice. If possible, please present UK-specific multipliers and include the option to use these in the model, if not already present.

The utility multipliers were based on the ICUROS study, and all countries included in the study (including UK). Unfortunately, UK-specific multipliers are currently not available from ICUROS. NICE's independent assessment group also used international ICUROS estimates in their recent assessment of non-bisphosphonates.[33]

B15. PRIORITY QUESTION: The ICUROS appears to include EQ-5D-3L data, EQ-VAS data and time trade-off (TTO) data. Please ensure that the multipliers included in the model are based only on EQ-5D-3L data.

The multipliers from the ICUROS study included in the model are only based on the EQ-5D-3L data.

B16. PRIORITY QUESTION: NICE TA464 (bisphosphonates for treating osteoporosis) also used utility multipliers from the ICUROS study, but the multipliers differ from those presented in the CS. Please explain the difference in values.

The hip fracture utility multipliers (first and subsequent years) are the same in the NICE's analysis and in UCB's analysis, albeit with more decimal points in UCB's analysis (0.55 vs 0.545 for first year and 0.86 vs 0.857 for subsequent years, respectively). Utility multipliers for vertebral

fractures were also similar (0.68 vs 0.671 for year 1 and 0.85 vs 0.841 for subsequent years in NICE's analysis and UCB's analysis, respectively). Utility multipliers for NHNV fractures were calculated for more types of fractures in UCB's analysis compared to those included in NICE's analysis. Detailed data from ICUROS on utilities for additional fracture types are included in the Appendix of a study by Kanis et al. (2018).[80]

The reason behind the differences in values is that UCB used utility multipliers from the ICUROS which had a larger sample size for fractured patients. The ICUROS estimates used in UCB's analysis had a total sample size for hip and vertebral fractures of about 3,000 patients, compared with about 1,000 patients in the publication that was used in NICE's analysis.[81]

B17. PRIORITY QUESTION: The CS states that the disutilities for multiple fractures are accounted for in a multiplicative approach. Please respond to the following points:

A. Was it possible for individuals to receive more than 1 acute multiplier at the same time?

Yes, it was possible for a patient to have a maximum two acute multipliers at the same time. This would happen if a patient experiences two fractures in the same year (i.e., in two consecutive cycles).

B. Did all patients enter the model with the full age-related general population utilities or were multipliers already applied to some patients?

No, QoL was adjusted so that patients with a fracture at the start of the model had QoL corresponding to the acute or chronic state of fracture (depending on time since fracture).

C. Please consider how plausible it is that multiple prior fractures have the same relative impact on HRQoL in the long-term (e.g. 5+ years after occurrence), when a new fracture is experienced in the last year.

A life-time impact on QoL after fracture is a common assumption in economic evaluations of osteoporosis treatment. In the recent assessment by NICE of non-bisphosphonates, the independent academic Assessment Group assumed that the quality-of-life multiplier was the same in the second year after fracture as in the subsequent years, with no restriction of duration of the impact.[33]

Many other economic evaluations have made the same assumptions, as identified by the systematic review in the aforementioned MTA.[33] Furthermore, the ESCEO/IOF guidelines recommends assuming QoL impact of fracture for all years after fracture, separated by an acute and a chronic multiplier.[59]

D. Please provide evidence that the included fracture types continue to affect HRQoL to the same extent 2 years, 5 years, 10 years and longer after occurrence. Please clarify that the model's assumptions regarding the length of time fractures are assumed to continue to

affect utility and consider the plausibility of these assumptions. Please add the option in the model to reduce the duration of impact of chronic (2nd year+) multipliers, if a lifetime impact of such fractures has been assumed.

A systematic review published in 2014 by Si et al. (2014) identified studies that assessed QoL after fracture for a follow-up at least 24 months.[82] Adachi et al. (2011) found that QoL (EQ-5D) remained lower 2 and 3 years after hip fractures compared with before the fracture (year 1 utility: 0.710, year 2: 0.720, year 3: 0.690).[83] Blomfeldt et al. (2005) showed that utility (EQ-5D) remained lower until year 5 after hip fracture, compared with before fracture. At year 1 the utility was 0.630, at year 2 0.640 and at year 5 0.620.[84] Ekström et al. (2009) showed that utility (EQ5D) remained lower after hip fracture until year 2 (year 1: 0.530, year 2: 0.520).[85]

Furthermore, since the approach of modelling chronic disutility for the remaining time horizon has been applied and accepted by NICE in a recent evaluation, this restriction has not been implemented in our model.[86]

E. Please add the option in the model to assume a maximum disutility approach (whereby only 1 multiplier is applied, for the most impactful fracture at any point in time) or any other approach or amendments to the multiplicative approach that the company considers could appropriately capture the impact of multiple fractures, both acute (in the last year) and chronic (second or more years).

UCB believe that the current multiplicative approach incorporated in the model remains the most appropriate assumption to calculate the impact of fractures on HRQoL. This approach to model QoL multiplicatively was validated by clinical experts in an internal economic advisory-board that was held in 2017. Furthermore, the IOF/ESCEO guidelines recommends to adjust QoL for multiple fractures.[59]

One alternative approach would be to adjust QoL only for the first two fractures that occur, however, this would have very minimal impact on the results, since few patients sustain more than 2 fractures. Please see Question B27 for these results.

B18. PRIORITY QUESTION: Page 43 of the CS states "[redacted]." Please provide the fracture utility decrements and multipliers which would be obtained from the ARCH HRQoL study and provide further justification as to why these are considered inappropriate.

Section 11.1 of the ARCH CSR details the HRQoL results from the trial. QoL estimates were collected at pre-determined discrete timepoints, which meant that any negative impact of a fracture on the QoL at a specific timepoint was diluted by patients who were fracture free at that timepoint. This rendered these estimates unsuitable for use in the model. A more detailed

explanation outlining the limitations of HRQoL collected in the ARCH trial can be found in the answer for Question A8 of this document.

The romosozumab cost-effectiveness model instead uses ICUROS as input reference for health-related utility (HRU) values because this study was specifically designed to assess the QoL impact of fractures on osteoporosis patients over time with the objective to allow the appropriate use of its findings in cost-effectiveness models. This is achieved in ICUROS by capturing the QoL impact of patients as soon as possible after a fracture occurs regardless of treatment, as opposed to the design of the ARCH study where QoL is assessed irrespective of fracture occurrence at predetermined discrete time points and always in relation to one of the treatments investigated during the trial. It is therefore not appropriate to use the QoL data collected in ARCH as input in a cost-effectiveness model because it does not provide robust HRU values which are sensitive to the decrease in QoL associated with fracture occurrence and does not provide treatment-unspecific HRU values which are needed for valid economic evaluations. Using HRU values from ARCH is therefore expected to underestimate the potential QoL gain with treatment. ICUROS was commissioned by the IOF and is the largest prospective study on osteoporosis quality by including over 7000 patients in 12 countries.[44, 45]

As highlighted in the company submission, the independent academic Assessment Group used ICUROS in NICE TA464 and intended to do so again in the suspended NICE MTA ID901.[33, 51] ICUROS has also previously been used as reference for HRU values in economic evaluations for romosozumab and denosumab to the Swedish Dental and Pharmaceutical Benefits Agency (TLV) in Sweden and for romosozumab to the Scottish Medicines Consortium (SmC) in Scotland, as well as other recently published cost-effectiveness studies.[87, 88] The use of ICUROS utilities is also recommended by ESCEO/IOF, with national ICUROS data if available, or otherwise the international version.[59]

B19. Please explain how the QALY loss of 0.0075 for gastrointestinal adverse events was calculated.

We applied the same assumption on disutility of GIAEs as the approach taken by the independent academic Assessment Group as part of TA464.[51]

A fixed decrement of 0.0075 per patient was added both in NICE's analyses and our model. The Assessment Report states that this disutility was based on a previous systematic review and economic evaluation by Stevenson et al.[89], however, the calculations that arrived at a disutility of 0.0075 are not described in these reports. It was decided to align with the approach taken Assessment Group nevertheless, due to a lack of other studies on the disutility of GIAEs.

Resource use and costs

B20. PRIORITY QUESTION: The analysis does not include administration costs for drugs that are administered via a subcutaneous injection, neither for romosozumab nor for the comparators in the scenario analyses. For romosozumab, the company justifies this by referring to their plans to set up a

Patient Support Program (PSP) that includes homecare service, an adherence support program, and training of injection techniques.

- **Please provide more details regarding these plans and specify the costs of services and health care resource to the NHS and PSS that when the PSP is in place would be borne by the company instead.**

The complete Evenity patient support programme consists of:

- Homecare delivery which is currently provided by two Homecare providers across the UK (Pharmaxo and Lloyds). There are four deliveries across the 12 month treatment period with Evenity.
- Retrieve device training from the IQVIA Nurse team once a patient receives the delivery of romosozumab. This remote training can be conducted either over the phone or using the "Attend anywhere" platform for video training. Face-by-face nurse training would be offered by UCB as an exception for patients who may be unable to self-inject (such as patients with reduced dexterity).
- Following the device training, the patient (in agreement with their prescribing physician) would be signed up to UCB's wider adherence programme for the duration of the 12 month treatment. The adherence programme consists of a mixture of phone calls, emails and SMS communications, with one of the final communications being a reminder to the patient to talk to their HCP about transitioning to an anti-resorptive therapy to maintain the BMD gains from their Evenity treatment.
- In addition to 15 minutes of nurse time associated with each subcutaneous injection of romosozumab, UCB's PSP could be expected to save the NHS at least one face to face clinic appointment for device training, as well as one or more face to face appointments to see how the patient is progressing. These costs are explored further in the below question.

- **Please provide the option in the model to include drug administration costs (i.e. for subcutaneous injections) that are borne by the NHS and PSS when the PSP is not in place for romosozumab, as well as for the relevant comparators that are used in scenario analyses.**

A scenario analysis has been conducted where administration cost of subcutaneous (SC) injections has been included for all relevant drugs. The cost (£9.5 per administration) is based on a 15-minute visit (based on £38 per hour for GP nurse contact time). PSSRU Unit Costs of Health and Social Care 2020 10.2 Nurse (GP practice). Unit costs available 2019/2020 based on 1,573 hours per year, which includes 225 working days minus sickness absence (8 days) and any training/study days as reported for all NHS staff groups.

In the scenario analysis, romosozumab is associated with 12 SC injections days (i.e. 24 injections) per year administered by a nurse; teriparatide 365 injections/year and denosumab 2 injections/year. The results are displayed in the table below.

The possibility to include cost for subcutaneous injections is available in the model. This can be added on sheet "Cost input". In the below scenario analyses, the "number of nurse visits" per year was changed to 12 for romosozumab, and 365 for teriparatide, and 2 for denosumab. Costs

of intravenous injections for zoledronate was already included in the original analyses of the CS. The inclusion of drug administration costs in these scenario analyses revealed to only have a minor impact on cost-effectiveness results, with romosozumab/alendronate remaining cost-effective against alendronate alone as was apparent in the base case of the CS.

Table 61: Scenario analyses for pairwise comparisons of romosozumab/alendronate versus other comparators including cost for subcutaneous administrations (PAS price for romosozumab)

Scenario	Technologies (ROMO/ALN versus)	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY) for ROMO/ALN versus comparator
-	ROMO (sc)/ALN (oral)	[redacted]	[redacted]	[redacted]				
1	Alendronate (oral)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£19,434
2	Teriparatide (Forsteo, sc) (24 months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Cost saving
3	Teriparatide (Forsteo, sc) (18 months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Cost saving
4	Teriparatide (Movymia, biosimilar, sc)/alendronate (oral)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Cost saving
5	Teriparatide (Forsteo, sc)/alendronate (oral)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Cost saving
6	Raloxifene (oral)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£396
7	Denosumab (sc)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£43,000
8	Risedronate (oral)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,953
9	Zoledronate (i.v.)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£21,129

Abbreviations: ICER: incremental cost-effectiveness ratio; i.v.: intravenous; LYG: life years gained; PAS: patient access scheme; QALYs: quality-adjusted life years; ROMO/ALN: romosozumab-to-alendronate; sc: subcutaneous.

B21. PRIORITY QUESTION: The costs during the first year following a fracture were sourced from Gutiérrez et al., 2011 for hip fractures and from Gutiérrez et al., 2012 for vertebral and non-hip-non-vertebral fractures. Gutiérrez et al., 2011 provide cost estimates both as total costs for patients who incurred a hip fracture as well as incremental costs of patients who incurred a hip fracture relative to matched controls. Since the estimates reported by Gutiérrez et al. pertain to the cost year 2006/2007, the costs were inflated to 2019/2020.

- **Please confirm that the total (i.e. not the incremental) cost estimates from Gutiérrez et al. were used in the analysis for patients who had a fracture but not for those who did not have a fracture, and please justify the appropriateness of this approach.**

The total costs based on Gutiérrez et al. (2011) were used in the analysis for patients who suffered a fracture. No medical costs were applied for those who did not suffer fracture.

The total costs rather than incremental costs were used in the model for two reasons. Firstly, both the incremental and total cost in the Gutiérrez et al. (2011) study are likely underestimated due to censoring bias. The follow-up time is shorter for the fractured cohort compared with the non-fractured cohort, which is likely due to higher mortality in the fractured cohort. This is not adjusted for in the two source papers.[90, 91] This underestimates costs but it is unknown to what extent. In the model, cost is applied for each cycle after the fracture (until the patient dies) and, as the cost input is unadjusted for censoring, the total costs would be underestimated in the model as well. Secondly, using total costs as opposed to incremental costs is the standard in economic evaluations, for example, in Jönsson, et al. (2011).[92]

- **Please include the option in the model to use either the total costs, whilst applying these to both patients with and without a fracture correspondingly, and the incremental costs of patients who had a fracture relative to those who did not, with the latter only applied to patients who had a fracture.**

- **Please provide details regarding which cost estimates were used and which indices were used to inflate the costs of fractures, to clarify exactly how the cost estimates used in the analysis were arrived at.**

The cost estimates for fractures, in their original price level and the inflated (2020 price level) cost estimates are described in Table 62 below. The inflation index to inflate the fracture costs are available in Table 63. The source for the inflation index is the Office for National Statistics dataset.[93]

Table 62: Fracture costs, in original price level and inflated to 2020 price level

Fracture type	First year cost, original estimate (£)	Source	Price year, original estimate	Inflated cost (£) 2020 price year
Hip fracture	9,936	Table 4 Gutierrez, L., et al.[91]	2007	13,203
Clinical vertebral fracture	2,180	Table 4 Gutierrez, L., et al.[91]	2007	2,897
NHNV fracture	1,604	Table 4 Gutierrez, L., et al.[91]	2007	2,131
	Subsequent years	**Source**	**Price year**	
Hip fracture	106	NICE Assessment Report: Bisphosphonates for preventing osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161)[51]	2015	115
Clinical vertebral fracture	332	NICE Assessment Report: Bisphosphonates for preventing osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161)[51]	2015	361

Abbreviations: NHNV: Non-hip non-vertebral

Table 63: Consumer price index, all items. 2015=100

Year, annual average	Index
2007	81.8
2008	84.7
2009	86.6
2010	89.4
2011	93.4
2012	96.1
2013	98.5
2014	100
2015	100
2016	100.7
2017	103.4
2018	105.9
2019	107.8

2020	108.7

Source: Office for National Statistics dataset, Table 20a, D7BT.[94]

- **Please justify the appropriateness of including rehabilitation costs only for hip fractures and not for other types of fractures. Please provide the option in the model to either include rehabilitation costs for all types of fractures for which these are relevant or exclude rehabilitation costs for all types of fractures.**

Hip fractures account for the majority of healthcare costs associated with osteoporosis, as well as having a major impact on patients' lives, resulting in increased morbidity, disability and mortality and diminished QoL compared to patients without a hip fracture. Based on the recent SCOPE study, women in the UK were considered to be at high risk (annual incidence of 405 in every 100,000) of experiencing a hip fracture compared to other European countries.[48]

Recent evidence has shown that anabolic agents have greater BMD gains compared to antiresorptive therapy, however, teriparatide has demonstrated little to no effect on fracture incidence and BMD at sites with a greater cortical bone component such as the hip and non-vertebral sites.[95-97]

There is therefore a clear unmet need for the availability of an effective treatment for hip fractures, such as romosozumab.

In addition, data on rehabilitation costs were not included for non-hip fractures as they were not available in the source data. Gutiérrez et al. (2012) estimated rehabilitation costs for hip fractures based on other studies, since they were not available in the THIN database, which is included in the cost estimate in the model.[91] Due to lack of evidence showing that non-hip fractures have an impact on rehabilitation costs, it was not been included in the model.

- **Please comment on the suitability of the hip fracture cost shown in Table 33 of the CS (£13,203), which is considerably higher than the cost used by the Assessment Group in NICE ID901 (£8,568; shown in Table 8 of the Assessment Report).**

NICE used the incremental cost of fracture (£8,568) which explains why the cost is higher (£13,2013) in our model. The justification for choosing total over incremental costs was provided above.

B22. PRIORITY QUESTION: Please provide details regarding how the annual drug and management costs that are listed in Table 31 of the CS were calculated.

Table 31 in the company submission shows the drug cost for the included treatments. The costs are sourced directly from the BNF database (BNF/NHS indicative price as described in the table) and no further calculations have been applied.

B23. PRIORITY QUESTION: Please explain whether the treatment costs as applied in the model are in line with treatment adherence as observed in the

treatment effectiveness results that are used to inform the model, and provide the option in the model to apply treatment costs in line with data on adherence (e.g. as provided in Table 25 in the CS) for all treatments considered in both base case and scenario analyses.

Treatment costs are only applied for as long as the patient is on treatment, therefore, treatment costs are "adjusted" for treatment adherence. No additional option has been added to the model.

B24. Please provide the rationale and functionality of the 'Morbidity cost shares' inputs on the 'Cost input' sheet of the model that is commented as an optional input.

The "morbidity cost share" are simply included as an optional input for when it is desirable to present costs depending on which ward/healthcare facility they occur. The numbers included do not have any impact on the calculation of total costs or the incremental cost-effectiveness ratio, but is only for presentation purposes. For this submission, the fracture related costs are presented as a total sum and not presented by setting.

B25. Please justify the appropriateness of assuming the costs of chemotherapy intravenous infusion for the administration of zoledronate.

In NICE TA464, the independent assessment group applied the reference cost for a day case delivery of a simple parenteral chemotherapy (HRG code: SB12Z at £245) to represent the cost of administration of zoledronate, as no alternative reference costs were identified which would cover day case admissions for the administration of a drug by infusion.[98]

The independent assessment group noted that "the outpatient cost for the same HRG code (SB12Z) was £165 suggesting that it is classification of this activity as a day case rather than the specific nature of chemotherapy that makes this more expensive than an outpatient endocrinology appointment."[98]

The independent assessment group therefore considered it reasonable to apply the day case reference cost for parenteral chemotherapy as a proxy for the cost of delivering zoledronate. Based on the same assumptions used in TA464, the cost for chemotherapy intravenous infusion was deemed a suitable proxy for the administration cost of zoledronate in this submission.

Cost effectiveness analyses

B26. PRIORITY QUESTION: Please provide a detailed explanation for the results of scenarios that demonstrate a large impact on the cost-effectiveness results when alternative values or assumptions are used, including start age and time horizon.

The parameters that had the largest impact on cost-effectiveness were time horizon, persistence, start age, and treatment effect on hip fractures. Below is a detailed description of the sensitivity analyses performed on each of these parameters.

- **Time horizon:** The base case time horizon was lifetime, meaning that the hypothetical patients were simulated from the start of treatment until death (or reached 100 years of

age, whichever came first). This is the standard time horizon for chronic diseases to account for the fact that treatments can continue to impact effects and costs throughout the patients' lifetime. Other, specific time horizons were tested as sensitivity analyses (5, 10, 15, and 20 years), where the patient was followed until the end of time horizon irrespective of whether the patient had died before that time point. Time horizons where patients are followed for a shorter time than the actual remaining lifetime (such as 5 or 10 years) increases the ICER. This is because fewer patients will have sustained fractures over the shorter time horizon, and consequently, the treatment will have avoided less fractures; thus, the effect on costs and QALYs is smaller. This impact of time horizon in osteoporosis models has been noted in previous economic evaluations, for example, a cost-effectiveness study of abaloparatide by Hiligsmann et al. (2020).[99]

- **Persistence:** As described in the CS, persistence is known to be sub-optimal for osteoporosis treatments. Changing persistence assumptions has a substantial impact on the cost-effectiveness results because lower/higher persistence reduces/improves the treatment effect since patients need to be on osteoporosis treatment for a while to allow treatment effectiveness to build-up. Lower/higher persistence is however offset by lower/higher treatment costs, and as such, the impact of persistence may be particularly large in treatments that have a low cost like alendronate. A study by Ström et al. (2009) described the importance of incorporating adherence in economic evaluation of osteoporosis models due to the potentially large effect on results.[100] The sensitivity of persistence assumptions has been demonstrated for example in a cost-effectiveness analysis of denosumab by Jönsson et al. (2011).[92]

- **Start age:** The start age in the base case is 74 years, which was the mean age of patients in the ARCH trial. Sensitivity analyses are included where the sensitivity to increasing and decreasing the age at which treatment is started is tested (50, 70, 60 and 80 years), keeping the other risk factors constant. These analyses demonstrate that the ICER is highest in the younger ages (50–60), slightly higher than the base case at age of 70, and lower than the base case at age 80. It should be noted that as only age is varied in these analyses, the fracture probability according to FRAX is lower in the age groups 50–70 compared with the base case, and higher than the base case in the age group 80. The ICER increases with start ages that are younger than the base case start age because with younger age, the patient has a lower fracture probability which is associated with poorer cost-effectiveness (since fewer fractures occur, and therefore fewer fractures can be prevented when being on treatment). An age older than the base case start age (80 years) slightly decreased the ICER, as older ages are associated with higher fracture probability. This relationship between age and cost-effectiveness is expected and has been shown in previous economic evaluations. For example, in the

study by Hiligsmann et al. (2020), the ICER of abaloparatide vs alendronate decreased from about $200,000 in a 50-year-old to about $70,000 in a 70-year-old.[99]

- **Treatment effect on hip fractures:** Assumptions related to the treatment effect on hip fractures have greater impact on cost-effectiveness compared with other fractures due to the large impact a hip fracture has on acute and long-term mortality, QoL and fracture-related costs (in particular the risk of moving to nursing home after the fracture). The large impact of changing hip fracture efficacy was also demonstrated in a Japanese cost-effectiveness study of denosumab where the ICER doubled when changing the relative risk of denosumab to the upper end of the confidence interval and nearly halved when changing it to the lower end of the confidence interval.[101]

Model validation

B27. Please provide a comparison of the distribution of fractures in the source data vs. the distribution of fractures in the simulation. The idea is to validate the statement on page 70 of the company submission "few patients experienced a third fracture in the source data".

In the source data, i.e., the Swedish real-world study of the risk of MOF in fractured patients, out of the 231,769 patients with at least one fracture, 7,656 patients (3.3%) had a third fracture over approximately 5.5 years of maximum follow-up data.[46] In the model simulation, 4.4% of patients had a third fracture over 5 years. These numbers are however not strictly comparable since in the source data, the first fracture could have happened at some point during the 5.5 years of follow-up, meaning that not all patients would have enough follow-up time to have developed a second or a third fracture.

Section C: Textual clarification and additional points

C1. Please correct the errors (#N/A and #NUM!) in the model 'PSA input' sheet.

The model without errors in the PSA input sheet has been sent alongside this response document (EVENITY CE Model_UK_2021-08-02).

C2. The macros included in the model are inside a password-protected VBA project.

A. Please provide the password for the VBA project.

While UCB had intended to submit the model with full access for the ERG's consideration and review, at present we are unfortunately not in a position that allows to share the password to the protected area of the economic model. This is because this area includes the FRAX algorithm, which is a third-party owned and patent protected resource that cannot be made accessible to further parties without the required legal contracts being put in place with NICE and/or ERG. UCB is currently engaging with NICE's Technology Appraisal Manager to explore legal options to be able to grant access to the password protected area of the economic model.

In an effort to be as transparent and responsive as possible to this question in the meanwhile, UCB has provided detailed descriptions (including screenshots) of all Macros in the password protected area of the economic model (except FRAX patent-protected ones) (Appendix C2B).

In addition, all VBA codes of the economic model (except FRAX patent-protected ones) are shared as .bas files in a zip folder (EVENITY CE Model VBA Modules). UCB is supportive to address any further inquiries from NICE and/or ERG relating to the password-protected area of the economic model (even after the clarification question stage), by e.g. providing a live "walk-through" of the password protected area (expect FRAX patent-protected areas) over a virtual or physical meeting with NICE and/or ERG.

B. Please provide a detailed explanation of the functionality and implementation for each macro included in the model.

Please see Appendix C2B for a detailed explanation of the functionality and implementation of each macro.

C3. PRIORITY QUESTION: Please include in the model 'Main settings' sheet the option to select all comparators included in the analyses.

Please see Appendix C3 for guidance on how to incorporate the comparators and the associated treatment sequences in the model.

References

1. National Osteoporosis Guidelines Group. NOGG 2017: Clinical Guideline for the Prevention and Treatment of Osteoporosis. Available at: https://www.sheffield.ac.uk/NOGG/NOGG%20Guideline%202017.pdf. Last accessed: 23 April 2021.
2. Pongchaiyakul C, Nguyen ND, Jones G, et al. Asymptomatic vertebral deformity as a major risk factor for subsequent fractures and mortality: a long-term prospective study. J Bone Miner Res 2005;20:1349-55.
3. FRAX - Fracture Risk Assessment Tool FAQ. Available at: https://www.sheffield.ac.uk/FRAX/faq.aspx. Last Accessed: 29th July 2021.
4. Pinedo-Villanueva R, Charokopou M, Toth E, et al. Imminent Fracture Risk Assessments in the UK FLS Setting: Implications and Challenges. Archives of Osteoporosis 2019;14:12.
5. Royal Osteoporosis Society (ROS). Effective Secondary Prevention of Fragility Fractures. Clinical Standards for Fracture Liaison Services, 2019.
6. British Orthopaedic Association Standards for Trauma. Fracture Liaison Services. Available at: https://www.boa.ac.uk/uploads/assets/41b78902-50b3-4197-86ee192efa68b4ab/233cee08-aef8-4604-b81132b119cf7351/fracture%20liaison%20services.pdf. Last accessed: 2nd August 2021.
7. Royal College of Physicians. Fracture Liaison Service Database Reporting. Available at: https://www.fffap.org.uk/FLS/charts.nsf/. Last accessed: 2nd August 2021.
8. McLellan AR, Wolowacz SE, Zimovetz EA, et al. Fracture liaison services for the evaluation and management of patients with osteoporotic fracture: a cost-effectiveness evaluation based on data collected over 8 years of service provision. Osteoporos Int 2011;22:2083-98.
9. Javaid MK, Sami A, Lems W, et al. A patient-level key performance indicator set to measure the effectiveness of fracture liaison services and guide quality improvement: a position paper of the IOF Capture the Fracture Working Group, National Osteoporosis Foundation and Fragility Fracture Network. Osteoporos Int 2020;31:1193-1204.
10. Osteoporotic fracture guideline. Available at https://www.guidelines.co.uk/musculoskeletal-and-joints-/osteoporotic-fracture-guideline/455546.article [accessed 08 June 2021].
11. Kanis JA, Johansson H, Harvey NC, et al. An assessment of intervention thresholds for very high fracture risk applied to the NOGG guidelines : A report for the National Osteoporosis Guideline Group (NOGG). Osteoporos Int 2021.
12. NHSX. "£36 million boost for AI technologies to revolutionise NHS care". Available at: https://www.nhsx.nhs.uk/news/36-million-boost-for-ai-technologies-to-revolutionise-nhs-care/. Accessed 26 July 2021.
13. National Institute for Health and Care Excellence. CG146. Osteoporosis: assessing the risk of fragility fracture. [accessed 23 April 2021] Available from: https://www.nice.org.uk/guidance/cg146/chapter/introduction.
14. Salkeld G, Cameron ID, Cumming RG, et al. Quality of life related to fear of falling and hip fracture in older women: a time trade off study. Bmj 2000;320:341-6.
15. Dyer SM, Crotty M, Fairhall N, et al. A Critical Review of the Long-Term Disability Outcomes Following Hip Fracture. BMC Geriatrics 2016;16:158.
16. Kanis JA, Oden A, Johnell O, et al. The Burden of Osteoporotic Fractures: A Method for Setting Intervention Thresholds. Osteoporos Int 2001;12:417–427.
17. Cooper C, Jakob F, Chinn C, et al. Fracture incidence and changes in quality of life in women with an inadequate clinical outcome from osteoporosis therapy: the

Observational Study of Severe Osteoporosis (OSSO). Osteoporos Int 2008;19:493-501.
18. Oglesby AK, Minshall ME, Shen W, et al. The impact of incident vertebral and non-vertebral fragility fractures on health-related quality of life in established postmenopausal osteoporosis: results from the teriparatide randomized, placebo-controlled trial in postmenopausal women. J Rheumatol 2003;30:1579-83.
19. Silverman SL, Minshall ME, Shen W, et al. The relationship of health-related quality of life to prevalent and incident vertebral fractures in postmenopausal women with osteoporosis: results from the Multiple Outcomes of Raloxifene Evaluation Study. Arthritis Rheum 2001;44:2611-9.
20. Osnes EK, Lofthus CM, Meyer HE, et al. Consequences of hip fracture on activities of daily life and residential needs. Osteoporos Int 2004;15:567-74.
21. Cauley JA, Thompson DE, Ensrud KC, et al. Risk of Mortality Following Clinical Fractures. Osteoporos Int 2000;11:556–561.
22. Center JR, Nguyen TV, Schneider D, et al. Mortality after all major types of osteoporotic fracture in men and women: an observational study. Lancet 1999;353:878-82.
23. Katsoulis M, Benetou V, Karapetyan T, et al. Excess mortality after hip fracture in elderly persons from Europe and the USA: the CHANCES project. J Intern Med 2017;281:300-310.
24. Vestergaard P, Rejnmark L, Mosekilde L. Increased mortality in patients with a hip fracture-effect of pre-morbid conditions and post-fracture complications. Osteoporos Int 2007;18:1583-93.
25. Ioannidis G, Papaioannou A, Hopman WM, et al. Relation between fractures and mortality: results from the Canadian Multicentre Osteoporosis Study. CMAJ 2009;181:265-71.
26. Svedbom A, Hernlund E, Ivergard M, et al. Osteoporosis in the European Union: A Compendium of Country-Specific Reports. Arch Osteoporos 2013;8:137.
27. Bliuc D, Nguyen ND, Nguyen TV, et al. Compound risk of high mortality following osteoporotic fracture and refracture in elderly women and men. J Bone Miner Res 2013;28:2317-24.
28. Bliuc D, Nguyen TV, Eisman JA, et al. The impact of nonhip nonvertebral fractures in elderly women and men. J Clin Endocrinol Metab 2014;99:415-23.
29. Gosch M, Druml T, Nicholas JA, et al. Fragility non-hip fracture patients are at risk. Arch Orthop Trauma Surg 2015;135:69-77.
30. Saag KG, Petersen J, Brandi ML, et al. Romosozumab or Alendronate for Fracture Prevention in Women with Osteoporosis. N Engl J Med 2017;377:1417–1427.
31. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Editors: Julian PT Higgins and Sally Green. Available at: https://handbook-5-1.cochrane.org/.
32. UCB Data on File. ARCH Clinical Study Report - July 2017.
33. National Institute for Health and Care Excellence (NICE). Assessment Group Report: Non-bisphosphonates for treating osteoporosis [ID901]. In: School of Health and Related Research (ScHARR) UoS, ed, 2019.
34. Simpson EL, Martyn-St James M, Hamilton J, et al. Clinical effectiveness of denosumab, raloxifene, romosozumab, and teriparatide for the prevention of osteoporotic fragility fractures: A systematic review and network meta-analysis. Bone 2020;130:115081.
35. Tsourdi E, Langdahl B, Cohen-Solal M, et al. Discontinuation of denosumab therapy for osteoporosis: a systematic review and position statement by ECTS. Bone 2017;105:11-17.
36. Miller PD, Bolognese MA, Lewiecki EM, et al. Effect of denosumab on bone density and turnover in postmenopausal women with low bone mass after long-term continued, discontinued, and restarting of therapy: a randomized blinded phase 2 clinical trial. Bone 2008;43:222-229.

37. Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Transactions on Modeling and Computer Simulation (TOMACS) 1998;8:3-30.
38. Stout NK, Goldie SJ. Keeping the noise down: common random numbers for disease simulation modeling. Health care management science 2008;11:399-406.
39. Johansson H, Siggeirsdóttir K, Harvey NC, et al. Imminent Risk of Fracture after Fracture. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2017;28:775–780.
40. Banefelt J, Åkesson K, Spångeus A, et al. Risk of imminent fracture following a previous fracture in a Swedish database study. Osteoporosis International 2019;30:601-609.
41. van Geel TA, van Helden S, Geusens PP, et al. Clinical Subsequent Fractures Cluster in Time After First Fractures. Ann Rheum Dis 2009;68:99–102.
42. Kanis JA, Cooper C, Rizzoli R, et al. Identification and management of patients at increased risk of osteoporotic fracture: outcomes of an ESCEO expert consensus meeting. Osteoporos Int 2017;28:2023-2034.
43. Kanis JA, Johansson H, Harvey NC, et al. Adjusting conventional FRAX estimates of fracture probability according to the recency of sentinel fractures. Osteoporosis International 2020;31:1817-1828.
44. Soreskog E, Borgstrom F, Lindberg I, et al. A novel economic framework to assess the cost-effectiveness of bone-forming agents in the prevention of fractures in patients with osteoporosis. Osteoporos Int 2021.
45. Soreskog E, Lindberg I, Kanis JA, et al. Cost-effectiveness of romosozumab for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture in Sweden. Osteoporos Int 2021.
46. Söreskog E, Ström O, Spångéus A, et al. Risk of major osteoporotic fracture after first, second and third fracture in Swedish women aged 50 years and older. Bone 2020;134:115286.
47. Kanis JA, Harvey NC, McCloskey E, et al. Algorithm for the management of patients at low, high and very high risk of osteoporotic fractures. Osteoporos Int 2020;31:1-12.
48. Kanis JA, Norton N, Harvey NC, et al. SCOPE 2021: a new scorecard for osteoporosis in Europe. Archives of Osteoporosis 2021;16:1-82.
49. Hernlund E, Svedbom A, Ivergård M, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos 2013;8:136.
50. Strom O, Jonsson B, Kanis JA. Intervention thresholds for denosumab in the UK using a FRAX(R)-based cost-effectiveness analysis. Osteoporos Int 2013;24:1491-502.
51. Davis S, James MMS, Sanderson J, et al. A systematic review and economic evaluation of bisphosphonates for the prevention of fragility fractures. Health Technology Assessment 2016;20:i-312.
52. Van Staa T, Dennison E, Leufkens Ha, et al. Epidemiology of fractures in England and Wales. Bone 2001;29:517-522.
53. Survey H. Available at http://healthsurvey.hscic.gov.uk/data-visualisation/data-visualisation/explore-the-trends/weight/adult/bmi.aspx.
54. Borgström F, Karlsson L, Ortsäter G, et al. Fragility fractures in Europe: burden, management and opportunities. Archives of Osteoporosis 2020;15:59.
55. van der Velde RY, Wyers CE, Curtis EM, et al. Secular Trends in Fracture Incidence in the UK Between 1990 And 2012. Osteoporos Int 2016;27:3197–3206.
56. Singer BR, McLauchlan GJ, Robinson CM, et al. Epidemiology of Fractures in 15,000 Adults: The Influence of Age and Gender. J Bone Joint Surg Br 1998;80:243–248.

57. Felsenberg D, Silman A, Lunt M, et al. Incidence of vertebral fracture in europe: results from the European Prospective Osteoporosis Study (EPOS). Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research 2002;17:716-724.
58. British Medical Journal - A glossary of EBM terms (Hazard Ratio). Available at: https://bestpractice.bmj.com/info/toolkit/ebm-toolbox/a-glossary-of-ebm-terms. Last accessed: 27th July 2021.
59. Hiligsmann M, Reginster JY, Tosteson ANA, et al. Recommendations for the conduct of economic evaluations in osteoporosis: outcomes of an experts' consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the US branch of the International Osteoporosis Foundation. Osteoporos Int 2019;30:45-57.
60. Li L, Roddam A, Gitlin M, et al. Persistence With Osteoporosis Medications Among Postmenopausal Women in the UK General Practice Research Database. Menopause 2012;19:33–40.
61. Landfeldt E, Ström O, Robbins S, et al. Adherence to treatment of primary osteoporosis and its association to fractures—the Swedish Adherence Register Analysis (SARA). Osteoporosis International 2012;23:433-443.
62. Office for National Statistics (ONS). Life Tables 2012-2014 Available at https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables [accessed 27 May 2021].
63. Strom O, Borgstrom F, Sen SS, et al. Cost-effectiveness of alendronate in the treatment of postmenopausal women in 9 European countries--an economic evaluation based on the fracture intervention trial. Osteoporos Int 2007;18:1047-61.
64. Borgstrom F, Johnell O, Kanis JA, et al. Cost Effectiveness of Raloxifene in the Treatment of Osteoporosis in Sweden: An Economic Evaluation Based on the MORE Study. Pharmacoeconomics 2004;22:1153–1165.
65. Parker M, Anand J. What is the true mortality of hip fractures? Public health 1991;105:443-446.
66. Kanis JA, Oden A, Johnell O, et al. The Components of Excess Mortality After Hip Fracture. Bone 2003;32:468–473.
67. Bliuc D, Nguyen ND, Milch VE, et al. Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. Jama 2009;301:513-521.
68. Vestergaard P, Rejnmark L, Mosekilde L. Osteoporosis is markedly underdiagnosed: a nationwide study from Denmark. Osteoporosis international 2005;16:134-141.
69. Van Staa T-P, Kanis JA, Geusens P, et al. The cost-effectiveness of bisphosphonates in postmenopausal women based on individual long-term fracture risks. Value in Health 2007;10:348-357.
70. Barrett JA, Baron JA, Beach ML. Mortality and Pulmonary Embolism after Fracture in the Elderly. Osteoporos Int 2003;14:889–894.
71. European Medicines Agency (EMA). EVENITY European Product Assessment Report. Available at https://www.ema.europa.eu/en/medicines/human/EPAR/evenity [accessed 26 July 2021].
72. Rothwell PM, Coull A, Giles M, et al. Change in stroke incidence, mortality, case-fatality, severity, and risk factors in Oxfordshire, UK from 1981 to 2004 (Oxford Vascular Study). The Lancet 2004;363:1925-1933.
73. Kannel WB, Gordon T. The Framingham Study: an epidemiological investigation of cardiovascular disease: Department of Health, Education, and Welfare, Public Health Service …, 1968.
74. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-1847.

75. Silverstein MD, Heit JA, Mohr DN, et al. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. Archives of internal medicine 1998;158:585-593.
76. Lindgren P, Lindström J, Tuomilehto J, et al. Lifestyle intervention to prevent diabetes in men and women with impaired glucose tolerance is cost-effective. International journal of technology assessment in health care 2007;23:177-183.
77. Ryder S, Fox K, Rane P, et al. A Systematic Review of Direct Cardiovascular Event Costs: An International Perspective. Pharmacoeconomics 2019;37:895-919.
78. Danese MD, Gleeson M, Kutikova L, et al. Estimating the economic burden of cardiovascular events in patients receiving lipid-modifying therapy in the UK. BMJ Open 2016;6:e011805.
79. Chapman RH, Liu LZ, Girase PG, et al. Determining initial and follow-up costs of cardiovascular events in a US managed care population. BMC cardiovascular disorders 2011;11:1-10.
80. Kanis JA, Johansson H, Oden A, et al. Characteristics of Recurrent Fractures. Osteoporos Int 2018;29:1747–1757.
81. Svedbom A, Borgström F, Hernlund E, et al. Quality of life after hip, vertebral, and distal forearm fragility fractures measured using the EQ-5D-3L, EQ-VAS, and time-trade-off: results from the ICUROS. Qual Life Res 2018;27:707-716.
82. Si L, Winzenberg TM, de Graaff B, et al. A systematic review and meta-analysis of utility-based quality of life for osteoporosis-related conditions. Osteoporos Int 2014;25:1987-97.
83. Adachi JD, Lyles KW, Colón-Emeric CS, et al. Zoledronic acid results in better health-related quality of life following hip fracture: the HORIZON-Recurrent Fracture Trial. Osteoporos Int 2011;22:2539-49.
84. Blomfeldt R, Törnkvist H, Ponzer S, et al. Comparison of internal fixation with total hip replacement for displaced femoral neck fractures. Randomized, controlled trial performed at four years. J Bone Joint Surg Am 2005;87:1680-8.
85. Ekström W, Németh G, Samnegård E, et al. Quality of life after a subtrochanteric fracture: a prospective cohort study on 87 elderly patients. Injury 2009;40:371-6.
86. Davis S, Simpson E, Martyn-St James M, Hamilton J, Rawdin A, Wong R, Goka E, Gittoes N, Selby P. A systematic review and economic evaluation of non-bisphosphonates for the prevention of osteoporotic fragility fractures (ID901). School of Health and Related Research (ScHARR), University of Sheffield, 2018. Available at: https://www.nice.org.uk/guidance/gid-ta10072/documents/assessment-report. Last Accessed: 29th July 2021.
87. Scottish Medicines Consortium. Medicines advice romosozumab (Evenity), 2020.
88. Tandvårds- och läkemedelsförmånsverket (The Dental and Pharmaceutical Benefits Agency). Underlag för beslut om subvention - Nyansökan [Decision for reimbursement], 2020.
89. Stevenson M, Jones ML, De Nigris E, 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;9:1-160.
90. Gutiérrez L, Roskell N, Castellsague J, et al. Study of the Incremental Cost and Clinical Burden of Hip Fractures in Postmenopausal Women in the United Kingdom. Journal of Medical Economics 2011;14:99–107.
91. Gutiérrez L, Roskell N, Castellsague J, et al. Clinical Burden and Incremental Cost of Fractures in Postmenopausal Women in the United Kingdom. Bone 2012;51:324–331.
92. Jonsson B, Strom O, Eisman JA, et al. Cost-Effectiveness of Denosumab for the Treatment of Postmenopausal Osteoporosis. Osteoporos Int 2011;22:967–982.
93. Office for National Statistics (ONS). Inflation Index. Table 20a, D7BT. Available at https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation). [accessed 26 July 2021].

94. Office for National Statistics. Consumer price inflation tables. 2021. Available at: https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation. Last Accessed: 27th July 2021.
95. Forsteo (teriparatide) SmPC. Available at: https://www.medicines.org.uk/emc/product/2215. Last accessed: 23 April 2021.
96. Oswald AJ, Berg K, Ralston SH, et al. Long-term effects of teriparatide followed by antiresorptive therapy on clinical outcomes in patients with severe spinal osteoporosis. Calcified tissue international 2019;105:148-155.
97. Langdahl BL, Libanati C, Crittenden DB, et al. Romosozumab (Sclerostin Monoclonal Antibody) Versus Teriparatide in Postmenopausal Women with Osteoporosis Transitioning from Oral Bisphosphonate Therapy: A Randomised, Open-Label, Phase 3 Trial. Lancet 2017;390:1585–1594.
98. National Institute for Health and Care Excellence (NICE). Bisphosphonates for treating osteoporosis [TA464]. Available at https://www.nice.org.uk/guidance/ta464 [accessed 22 April 2021], 2019.
99. Hiligsmann M, Williams SA, Fitzpatrick LA, et al. Cost-effectiveness of sequential treatment with abaloparatide followed by alendronate vs. alendronate monotherapy in women at increased risk of fracture: A US payer perspective. Semin Arthritis Rheum 2020;50:394-400.
100. Ström O, Borgström F, Kanis JA, et al. Incorporating adherence into health economic modelling of osteoporosis. Osteoporos Int 2009;20:23-34.
101. Yoshizawa T, Nishino T, Okubo I, et al. Cost-effectiveness analysis of drugs for osteoporosis treatment in elderly Japanese women at high risk of fragility fractures: comparison of denosumab and weekly alendronate. Arch Osteoporos 2018;13:94.

Appendices

Appendix A1

The full details of the second update searches are detailed in Table 64.

Table 64: Summary of the PubMed searches for the second SLR update

No.	PubMed Search	Search hits
#1	((" romosozumab"[tiab] OR " romosozumab"[tt] OR " romosozumab"[mh] OR " romosozumab"[rn] OR " romosozumab"[nm]) OR ("Evenity"[tiab] OR "Evenity"[tt] OR "Evenity"[mh] OR "Evenity"[rn] OR "Evenity"[nm]) OR ("AMG785"[tiab] OR "AMG785"[tt] OR "AMG785"[mh] OR "AMG785"[rn] OR "AMG785"[nm]) OR ("AMG-785"[tiab] OR "AMG-785"[tt] OR "AMG-785"[mh] OR "AMG-785"[rn] OR "AMG-785"[nm]) OR ("cdp-7851"[tiab] OR "cdp-7851"[tt] OR "cdp-7851"[mh] OR "cdp-7851"[rn] OR "cdp-7851"[nm]) OR ("cdp7851"[tiab] OR "cdp7851"[tt] OR "cdp7851"[mh] OR "cdp7851"[rn] OR "cdp7851"[nm]) OR ("909395-70-6"[tiab] OR "909395-70-6"[tt] OR "909395-70-6"[mh] OR "909395-70-6"[rn] OR "909395-70-6"[nm]))	225
#2	"Teriparatide"[mesh:noexp]	1944
#3	(Teriparatide or Forteo or Forsteo or chs-13340 or chs13340 or ly-333334 or ly333334 or parathar or "parathormone 1 34"[mesh] or "parathyroid hormone 1-34"[mesh] or "pth[1-34]" or "sun-e3001"[mesh] or "sune3001"[mesh] or 12583-68-5 or 52232-67-4)[tiab]	3828
#4	#2 OR #3	3828
#5	"Alendronate"[mesh:noexp]	3676
#6	(("alendronic acid"[tiab] OR "alendronic acid"[tt] OR "alendronic acid"[mh] OR "alendronic acid"[rn] OR "alendronic acid"[nm]) OR ("alendronate"[tiab] OR "alendronate"[tt] OR "alendronate"[mh] OR "alendronate"[rn] OR "alendronate"[nm]) OR ("alenato"[tiab] OR "alenato"[tt] OR "alenato"[mh] OR "alenato"[rn] OR "alenato"[nm]) OR ("alend"[tiab] OR "alend"[tt] OR "alend"[mh] OR "alend"[rn] OR "alend"[nm]) OR ("alendros"[tiab] OR "alendros"[tt] OR "alendros"[mh] OR "alendros"[rn] OR "alendros"[nm]) OR ("alovell"[tiab] OR "alovell"[tt] OR "alovell"[mh] OR "alovell"[rn] OR "alovell"[nm]) OR ("arendal"[tiab] OR "arendal"[tt] OR "arendal"[mh] OR "arendal"[rn] OR "arendal"[nm]) OR ("bifemelan"[tiab] OR "bifemelan"[tt] OR "bifemelan"[mh] OR "bifemelan"[rn] OR "bifemelan"[nm]) OR ("bifosa"[tiab] OR "bifosa"[tt] OR "bifosa"[mh] OR "bifosa"[rn] OR "bifosa"[nm]) OR ("binosto"[tiab] OR "binosto"[tt] OR "binosto"[mh] OR "binosto"[rn] OR "binosto"[nm]) OR ("bonapex"[tiab] OR "bonapex"[tt] OR "bonapex"[mh] OR "bonapex"[rn] OR "bonapex"[nm]) OR ("defixal"[tiab] OR "defixal"[tt] OR "defixal"[mh] OR "defixal"[rn] OR "defixal"[nm]) OR ("dronal"[tiab] OR "dronal"[tt] OR "dronal"[mh] OR "dronal"[rn] OR "dronal"[nm]) OR ("endronax"[tiab] OR "endronax"[tt] OR "endronax"[mh] OR "endronax"[rn] OR "endronax"[nm]) OR ("eucalen"[tiab] OR "eucalen"[tt] OR "eucalen"[mh] OR "eucalen"[rn] OR "eucalen"[nm]) OR ("fixopan"[tiab] OR "fixopan"[tt] OR "fixopan"[mh] OR "fixopan"[rn] OR "fixopan"[nm]) OR ("fosalan"[tiab] OR "fosalan"[tt] OR "fosalan"[mh] OR "fosalan"[rn] OR "fosalan"[nm]) OR ("fosamax"[tiab] OR "fosamax"[tt] OR "fosamax"[mh] OR "fosamax"[rn] OR "fosamax"[nm]) OR ("fosmin"[tiab] OR "fosmin"[tt] OR "fosmin"[mh] OR "fosmin"[rn] OR "fosmin"[nm]) OR ("fosval"[tiab] OR "fosval"[tt] OR "fosval"[mh] OR "fosval"[rn] OR "fosval"[nm]) OR ("marvil"[tiab] OR "marvil"[tt] OR "marvil"[mh] OR "marvil"[rn] OR "marvil"[nm]) OR ("maxibone"[tiab] OR "maxibone"[tt] OR "maxibone"[mh] OR "maxibone"[rn] OR "maxibone"[nm]) OR ("mk-0217"[tiab] OR "mk-0217"[tt] OR "mk-0217"[mh] OR "mk-0217"[rn] OR "mk-0217"[nm]) OR ("mk-217"[tiab] OR "mk-217"[tt] OR "mk-217"[mh] OR	5650

	"mk-217"[rn] OR "mk-217"[nm]) OR ("mk0217"[tiab] OR "mk0217"[tt] OR "mk0217"[mh] OR "mk0217"[rn] OR "mk0217"[nm]) OR ("mk217"[tiab] OR "mk217"[tt] OR "mk217"[mh] OR "mk217"[rn] OR "mk217"[nm]) OR ("neobon"[tiab] OR "neobon"[tt] OR "neobon"[mh] OR "neobon"[rn] OR "neobon"[nm]) OR ("oncalst"[tiab] OR "oncalst"[tt] OR "oncalst"[mh] OR "oncalst"[rn] OR "oncalst"[nm]) OR ("onclast"[tiab] OR "onclast"[tt] OR "onclast"[mh] OR "onclast"[rn] OR "onclast"[nm]) OR ("osdron"[tiab] OR "osdron"[tt] OR "osdron"[mh] OR "osdron"[rn] OR "osdron"[nm]) OR ("osdronat"[tiab] OR "osdronat"[tt] OR "osdronat"[mh] OR "osdronat"[rn] OR "osdronat"[nm]) OR ("oseotenk"[tiab] OR "oseotenk"[tt] OR "oseotenk"[mh] OR "oseotenk"[rn] OR "oseotenk"[nm]) OR ("osficar"[tiab] OR "osficar"[tt] OR "osficar"[mh] OR "osficar"[rn] OR "osficar"[nm]) OR ("oslene"[tiab] OR "oslene"[tt] OR "oslene"[mh] OR "oslene"[rn] OR "oslene"[nm]) OR ("osteofar"[tiab] OR "osteofar"[tt] OR "osteofar"[mh] OR "osteofar"[rn] OR "osteofar"[nm]) OR ("osteofos"[tiab] OR "osteofos"[tt] OR "osteofos"[mh] OR "osteofos"[rn] OR "osteofos"[nm]) OR ("osteopor"[tiab] OR "osteopor"[tt] OR "osteopor"[mh] OR "osteopor"[rn] OR "osteopor"[nm]) OR ("osteosan"[tiab] OR "osteosan"[tt] OR "osteosan"[mh] OR "osteosan"[rn] OR "osteosan"[nm]) OR ("osteovan"[tiab] OR "osteovan"[tt] OR "osteovan"[mh] OR "osteovan"[rn] OR "osteovan"[nm]) OR ("osticalcin"[tiab] OR "osticalcin"[tt] OR "osticalcin"[mh] OR "osticalcin"[rn] OR "osticalcin"[nm]) OR ("porosal"[tiab] OR "porosal"[tt] OR "porosal"[mh] OR "porosal"[rn] OR "porosal"[nm]) OR ("teiroc"[tiab] OR "teiroc"[tt] OR "teiroc"[mh] OR "teiroc"[rn] OR "teiroc"[nm]) OR ("tibolene"[tiab] OR "tibolene"[tt] OR "tibolene"[mh] OR "tibolene"[rn] OR "tibolene"[nm]) OR ("voroste"[tiab] OR "voroste"[tt] OR "voroste"[mh] OR "voroste"[rn] OR "voroste"[nm]) OR ("Fosavance"[tiab] OR "Fosavance"[tt] OR "Fosavance"[mh] OR "Fosavance"[rn] OR "Fosavance"[nm]) OR ("Adrovance"[tiab] OR "Adrovance"[tt] OR "Adrovance"[mh] OR "Adrovance"[rn] OR "Adrovance"[nm]) OR ("Vantavo"[tiab] OR "Vantavo"[tt] OR "Vantavo"[mh] OR "Vantavo"[rn] OR "Vantavo"[nm]) OR ("Binosto"[tiab] OR "Binosto"[tt] OR "Binosto"[mh] OR "Binosto"[rn] OR "Binosto"[nm]) OR ("mylan"[tiab] OR "mylan"[tt] OR "mylan"[mh] OR "mylan"[rn] OR "mylan"[nm]) OR ("Adronat"[tiab] OR "Adronat"[tt] OR "Adronat"[mh] OR "Adronat"[rn] OR "Adronat"[nm]) OR ("Alendro"[tiab] OR "Alendro"[tt] OR "Alendro"[mh] OR "Alendro"[rn] OR "Alendro"[nm]) OR ("Alendraccord"[tiab] OR "Alendraccord"[tt] OR "Alendraccord"[mh] OR "Alendraccord"[rn] OR "Alendraccord"[nm]) OR ("Alendrobell"[tiab] OR "Alendrobell"[tt] OR "Alendrobell"[mh] OR "Alendrobell"[rn] OR "Alendrobell"[nm]) OR ("Alendrocor-10"[tiab] OR "Alendrocor-10"[tt] OR "Alendrocor-10"[mh] OR "Alendrocor-10"[rn] OR "Alendrocor-10"[nm]) OR ("Densate-70"[tiab] OR "Densate-70"[tt] OR "Densate-70"[mh] OR "Densate-70"[rn] OR "Densate-70"[nm]) OR ("Dronalen-Plus"[tiab] OR "Dronalen-Plus"[tt] OR "Dronalen-Plus"[mh] OR "Dronalen-Plus"[rn] OR "Dronalen-Plus"[nm]) OR ("Ossmax"[tiab] OR "Ossmax"[tt] OR "Ossmax"[mh] OR "Ossmax"[rn] OR "Ossmax"[nm]) OR ("66376-36-1"[tiab] OR "66376-36-1"[tt] OR "66376-36-1"[mh] OR "66376-36-1"[rn] OR "66376-36-1"[nm]))	
#7	#5 OR #6	5650
#8	Risedronate Sodium[mesh]	1183
#9	(("risedronic acid"[tiab] OR "risedronic acid"[tt] OR "risedronic acid"[mh] OR "risedronic acid"[rn] OR "risedronic acid"[nm]) OR ("actonel"[tiab] OR "actonel"[tt] OR "actonel"[mh] OR "actonel"[rn] OR "actonel"[nm]) OR ("atelvia"[tiab] OR "atelvia"[tt] OR "atelvia"[mh] OR "atelvia"[rn] OR "atelvia"[nm]) OR ("benet"[tiab] OR "benet"[tt] OR "benet"[mh] OR "benet"[rn] OR "benet"[nm]) OR ("ne-58095"[tiab] OR "ne-58095"[tt] OR "ne-58095"[mh] OR "ne-58095"[rn] OR "ne-58095"[nm]) OR ("ne58095"[tiab] OR "ne58095"[tt] OR "ne58095"[mh] OR "ne58095"[rn] OR "ne58095"[nm]) OR ("optinate"[tiab] OR "optinate"[tt] OR "optinate"[mh] OR "optinate"[rn] OR "optinate"[nm]) OR ("ribastamin"[tiab] OR "ribastamin"[tt] OR "ribastamin"[mh] OR "ribastamin"[rn] OR "ribastamin"[nm]) OR ("risedronate"[tiab] OR "risedronate"[tt] OR	2236

	"risedronate"[mh] OR "risedronate"[rn] OR "risedronate"[nm]) OR ("Acris"[tiab] OR "Acris"[tt] OR "Acris"[mh] OR "Acris"[rn] OR "Acris"[nm]) OR ("Risedro"[tiab] OR "Risedro"[tt] OR "Risedro"[mh] OR "Risedro"[rn] OR "Risedro"[nm]) OR ("benet"[tiab] OR "benet"[tt] OR "benet"[mh] OR "benet"[rn] OR "benet"[nm]) OR ("CO Risedrocal Combo Kit"[tiab] OR "CO Risedrocal Combo Kit"[tt] OR "CO Risedrocal Combo Kit"[mh] OR "CO Risedrocal Combo Kit"[rn] OR "CO Risedrocal Combo Kit"[nm]) OR ("aktonate"[tiab] OR "aktonate"[tt] OR "aktonate"[mh] OR "aktonate"[rn] OR "aktonate"[nm]) OR ("bonna"[tiab] OR "bonna"[tt] OR "bonna"[mh] OR "bonna"[rn] OR "bonna"[nm]) OR ("cladronate"[tiab] OR "cladronate"[tt] OR "cladronate"[mh] OR "cladronate"[rn] OR "cladronate"[nm]) OR ("ductonar"[tiab] OR "ductonar"[tt] OR "ductonar"[mh] OR "ductonar"[rn] OR "ductonar"[nm]) OR ("goyart"[tiab] OR "goyart"[tt] OR "goyart"[mh] OR "goyart"[rn] OR "goyart"[nm]) OR ("melenor"[tiab] OR "melenor"[tt] OR "melenor"[mh] OR "melenor"[rn] OR "melenor"[nm]) OR ("ostenel"[tiab] OR "ostenel"[tt] OR "ostenel"[mh] OR "ostenel"[rn] OR "ostenel"[nm]) OR ("osteodronate"[tiab] OR "osteodronate"[tt] OR "osteodronate"[mh] OR "osteodronate"[rn] OR "osteodronate"[nm]) OR ("ribastamin duo rigat"[tiab] OR "ribastamin duo rigat"[tt] OR "ribastamin duo rigat"[mh] OR "ribastamin duo rigat"[rn] OR "ribastamin duo rigat"[nm]) OR ("risate"[tiab] OR "risate"[tt] OR "risate"[mh] OR "risate"[rn] OR "risate"[nm]) OR ("risedron"[tiab] OR "risedron"[tt] OR "risedron"[mh] OR "risedron"[rn] OR "risedron"[nm]) OR ("risedrogen"[tiab] OR "risedrogen"[tt] OR "risedrogen"[mh] OR "risedrogen"[rn] OR "risedrogen"[nm]) OR ("risendronat"[tiab] OR "risendronat"[tt] OR "risendronat"[mh] OR "risendronat"[rn] OR "risendronat"[nm]) OR ("risemylan"[tiab] OR "risemylan"[tt] OR "risemylan"[mh] OR "risemylan"[rn] OR "risemylan"[nm]) OR ("risendal"[tiab] OR "risendal"[tt] OR "risendal"[mh] OR "risendal"[rn] OR "risendal"[nm]) OR ("isendros"[tiab] OR "isendros"[tt] OR "isendros"[mh] OR "isendros"[rn] OR "isendros"[nm]) OR ("risetab"[tiab] OR "risetab"[tt] OR "risetab"[mh] OR "risetab"[rn] OR "risetab"[nm]) OR ("risofos"[tiab] OR "risofos"[tt] OR "risofos"[mh] OR "risofos"[rn] OR "risofos"[nm]) OR ("risonato"[tiab] OR "risonato"[tt] OR "risonato"[mh] OR "risonato"[rn] OR "risonato"[nm]) OR ("salost"[tiab] OR "salost"[tt] OR "salost"[mh] OR "salost"[rn] OR "salost"[nm]) OR ("tracost"[tiab] OR "tracost"[tt] OR "tracost"[mh] OR "tracost"[rn] OR "tracost"[nm]) OR ("acrel"[tiab] OR "acrel"[tt] OR "acrel"[mh] OR "acrel"[rn] OR "acrel"[nm]) OR ("actomax"[tiab] OR "actomax"[tt] OR "actomax"[mh] OR "actomax"[rn] OR "actomax"[nm]) OR ("actojenic"[tiab] OR "actojenic"[tt] OR "actojenic"[mh] OR "actojenic"[rn] OR "actojenic"[nm]) OR ("actokit"[tiab] OR "actokit"[tt] OR "actokit"[mh] OR "actokit"[rn] OR "actokit"[nm]) OR ("arilex"[tiab] OR "arilex"[tt] OR "arilex"[mh] OR "arilex"[rn] OR "arilex"[nm]) OR ("atconate"[tiab] OR "atconate"[tt] OR "atconate"[mh] OR "atconate"[rn] OR "atconate"[nm]) OR ("bondapen"[tiab] OR "bondapen"[tt] OR "bondapen"[mh] OR "bondapen"[rn] OR "bondapen"[nm]) OR ("boneact"[tiab] OR "boneact"[tt] OR "boneact"[mh] OR "boneact"[rn] OR "boneact"[nm]) OR ("boncur"[tiab] OR "boncur"[tt] OR "boncur"[mh] OR "boncur"[rn] OR "boncur"[nm]) OR ("bonmate"[tiab] OR "bonmate"[tt] OR "bonmate"[mh] OR "bonmate"[rn] OR "bonmate"[nm]) OR ("bontonel"[tiab] OR "bontonel"[tt] OR "bontonel"[mh] OR "bontonel"[rn] OR "bontonel"[nm]) OR ("bontrol"[tiab] OR "bontrol"[tt] OR "bontrol"[mh] OR "bontrol"[rn] OR "bontrol"[nm]) OR ("claronate"[tiab] OR "claronate"[tt] OR "claronate"[mh] OR "claronate"[rn] OR "claronate"[nm]) OR ("enospag"[tiab] OR "enospag"[tt] OR "enospag"[mh] OR "enospag"[rn] OR "enospag"[nm]) OR ("fodren"[tiab] OR "fodren"[tt] OR "fodren"[mh] OR "fodren"[rn] OR "fodren"[nm]) OR ("juverital"[tiab] OR "juverital"[tt] OR "juverital"[mh] OR "juverital"[rn] OR "juverital"[nm]) OR ("medeoros"[tiab] OR "medeoros"[tt] OR "medeoros"[mh] OR "medeoros"[rn] OR "medeoros"[nm]) OR ("miosen"[tiab] OR "miosen"[tt] OR "miosen"[mh] OR "miosen"[rn] OR "miosen"[nm]) OR ("natalox"[tiab] OR "natalox"[tt] OR "natalox"[mh] OR "natalox"[rn] OR "natalox"[nm]) OR ("norifax"[tiab] OR "norifax"[tt] OR "norifax"[mh] OR "norifax"[rn] OR "norifax"[nm]) OR	

	("norsed"[tiab] OR "norsed"[tt] OR "norsed"[mh] OR "norsed"[rn] OR "norsed"[nm]) OR ("osodens"[tiab] OR "osodens"[tt] OR "osodens"[mh] OR "osodens"[rn] OR "osodens"[nm]) OR ("osteoron"[tiab] OR "osteoron"[tt] OR "osteoron"[mh] OR "osteoron"[rn] OR "osteoron"[nm]) OR ("ostron"[tiab] OR "ostron"[tt] OR "ostron"[mh] OR "ostron"[rn] OR "ostron"[nm]) OR ("pexalit"[tiab] OR "pexalit"[tt] OR "pexalit"[mh] OR "pexalit"[rn] OR "pexalit"[nm]) OR ("tentop"[tiab] OR "tentop"[tt] OR "tentop"[mh] OR "tentop"[rn] OR "tentop"[nm]) OR ("resorpate"[tiab] OR "resorpate"[tt] OR "resorpate"[mh] OR "resorpate"[rn] OR "resorpate"[nm]) OR ("retonel"[tiab] OR "retonel"[tt] OR "retonel"[mh] OR "retonel"[rn] OR "retonel"[nm]) OR ("ribastamin"[tiab] OR "ribastamin"[tt] OR "ribastamin"[mh] OR "ribastamin"[rn] OR "ribastamin"[nm]) OR ("ribidron"[tiab] OR "ribidron"[tt] OR "ribidron"[mh] OR "ribidron"[rn] OR "ribidron"[nm]) OR ("ribone"[tiab] OR "ribone"[tt] OR "ribone"[mh] OR "ribone"[rn] OR "ribone"[nm]) OR ("richbone"[tiab] OR "richbone"[tt] OR "richbone"[mh] OR "richbone"[rn] OR "richbone"[nm]) OR ("ridbone"[tiab] OR "ridbone"[tt] OR "ridbone"[mh] OR "ridbone"[rn] OR "ridbone"[nm]) OR ("ridron"[tiab] OR "ridron"[tt] OR "ridron"[mh] OR "ridron"[rn] OR "ridron"[nm]) OR ("ridrone"[tiab] OR "ridrone"[tt] OR "ridrone"[mh] OR "ridrone"[rn] OR "ridrone"[nm]) OR ("risadican"[tiab] OR "risadican"[tt] OR "risadican"[mh] OR "risadican"[rn] OR "risadican"[nm]) OR ("risbon"[tiab] OR "risbon"[tt] OR "risbon"[mh] OR "risbon"[rn] OR "risbon"[nm]) OR ("risebon"[tiab] OR "risebon"[tt] OR "risebon"[mh] OR "risebon"[rn] OR "risebon"[nm]) OR ("risebone"[tiab] OR "risebone"[tt] OR "risebone"[mh] OR "risebone"[rn] OR "risebone"[nm]) OR ("risedon"[tiab] OR "risedon"[tt] OR "risedon"[mh] OR "risedon"[rn] OR "risedon"[nm]) OR ("risedreenos"[tiab] OR "risedreenos"[tt] OR "risedreenos"[mh] OR "risedreenos"[rn] OR "risedreenos"[nm]) OR ("risedronaat"[tiab] OR "risedronaat"[tt] OR "risedronaat"[mh] OR "risedronaat"[rn] OR "risedronaat"[nm]) OR ("riselib"[tiab] OR "riselib"[tt] OR "riselib"[mh] OR "riselib"[rn] OR "riselib"[nm]) OR ("risemed"[tiab] OR "risemed"[tt] OR "risemed"[mh] OR "risemed"[rn] OR "risemed"[nm]) OR ("risedrenos"[tiab] OR "risedrenos"[tt] OR "risedrenos"[mh] OR "risedrenos"[rn] OR "risedrenos"[nm]) OR ("risenex"[tiab] OR "risenex"[tt] OR "risenex"[mh] OR "risenex"[rn] OR "risenex"[nm]) OR ("risenil"[tiab] OR "risenil"[tt] OR "risenil"[mh] OR "risenil"[rn] OR "risenil"[nm]) OR ("riseto"[tiab] OR "riseto"[tt] OR "riseto"[mh] OR "riseto"[rn] OR "riseto"[nm]) OR ("risetron"[tiab] OR "risetron"[tt] OR "risetron"[mh] OR "risetron"[rn] OR "risetron"[nm]) OR ("resmyl"[tiab] OR "resmyl"[tt] OR "resmyl"[mh] OR "resmyl"[rn] OR "resmyl"[nm]) OR ("risofos"[tiab] OR "risofos"[tt] OR "risofos"[mh] OR "risofos"[rn] OR "risofos"[nm]) OR ("risonate"[tiab] OR "risonate"[tt] OR "risonate"[mh] OR "risonate"[rn] OR "risonate"[nm]) OR ("risonato"[tiab] OR "risonato"[tt] OR "risonato"[mh] OR "risonato"[rn] OR "risonato"[nm]) OR ("risostad"[tiab] OR "risostad"[tt] OR "risostad"[mh] OR "risostad"[rn] OR "risostad"[nm]) OR ("ristonat"[tiab] OR "ristonat"[tt] OR "ristonat"[mh] OR "ristonat"[rn] OR "ristonat"[nm]) OR ("sedron"[tiab] OR "sedron"[tt] OR "sedron"[mh] OR "sedron"[rn] OR "sedron"[nm]) OR ("seralis"[tiab] OR "seralis"[tt] OR "seralis"[mh] OR "seralis"[rn] OR "seralis"[nm]) OR ("tecnodron"[tiab] OR "tecnodron"[tt] OR "tecnodron"[mh] OR "tecnodron"[rn] OR "tecnodron"[nm]) OR ("tevanel"[tiab] OR "tevanel"[tt] OR "tevanel"[mh] OR "tevanel"[rn] OR "tevanel"[nm]) OR ("varibona"[tiab] OR "varibona"[tt] OR "varibona"[mh] OR "varibona"[rn] OR "varibona"[nm]) OR ("norifaz"[tiab] OR "norifaz"[tt] OR "norifaz"[mh] OR "norifaz"[rn] OR "norifaz"[nm]) OR ("zectoel"[tiab] OR "zectoel"[tt] OR "zectoel"[mh] OR "zectoel"[rn] OR "zectoel"[nm]) OR ("acridon"[tiab] OR "acridon"[tt] OR "acridon"[mh] OR "acridon"[rn] OR "acridon"[nm]) OR ("ridroqueen"[tiab] OR "ridroqueen"[tt] OR "ridroqueen"[mh] OR "ridroqueen"[rn] OR "ridroqueen"[nm]) OR ("105462-24-6"[tiab] OR "105462-24-6"[tt] OR "105462-24-6"[mh] OR "105462-24-6"[rn] OR "105462-24-6"[nm]) OR ("122458-82-6"[tiab] OR "122458-82-6"[tt] OR "122458-82-6"[mh] OR "122458-82-6"[rn] OR "122458-82-6"[nm]))	
#10	#8 OR #9	2236

#11	(("ibandronate"[tiab] OR "ibandronate"[tt] OR "ibandronate"[mh] OR "ibandronate"[rn] OR "ibandronate"[nm]) OR ("ibandronic acid"[tiab] OR "ibandronic acid"[tt] OR "ibandronic acid"[mh] OR "ibandronic acid"[rn] OR "ibandronic acid"[nm]) OR ("bonviva"[tiab] OR "bonviva"[tt] OR "bonviva"[mh] OR "bonviva"[rn] OR "bonviva"[nm]) OR ("bondronat"[tiab] OR "bondronat"[tt] OR "bondronat"[mh] OR "bondronat"[rn] OR "bondronat"[nm]) OR ("bondronate"[tiab] OR "bondronate"[tt] OR "bondronate"[mh] OR "bondronate"[rn] OR "bondronate"[nm]) OR ("boniva"[tiab] OR "boniva"[tt] OR "boniva"[mh] OR "boniva"[rn] OR "boniva"[nm]) OR ("destara"[tiab] OR "destara"[tt] OR "destara"[mh] OR "destara"[rn] OR "destara"[nm]) OR ("bm-210955"[tiab] OR "bm-210955"[tt] OR "bm-210955"[mh] OR "bm-210955"[rn] OR "bm-210955"[nm]) OR ("bm210955"[tiab] OR "bm210955"[tt] OR "bm210955"[mh] OR "bm210955"[rn] OR "bm210955"[nm]) OR ("bondenza"[tiab] OR "bondenza"[tt] OR "bondenza"[mh] OR "bondenza"[rn] OR "bondenza"[nm]) OR ("iasibon"[tiab] OR "iasibon"[tt] OR "iasibon"[mh] OR "iasibon"[rn] OR "iasibon"[nm]) OR ("ibandronico"[tiab] OR "ibandronico"[tt] OR "ibandronico"[mh] OR "ibandronico"[rn] OR "ibandronico"[nm]) OR ("alvodron"[tiab] OR "alvodron"[tt] OR "alvodron"[mh] OR "alvodron"[rn] OR "alvodron"[nm]) OR ("alvodronic"[tiab] OR "alvodronic"[tt] OR "alvodronic"[mh] OR "alvodronic"[rn] OR "alvodronic"[nm]) OR ("bandro"[tiab] OR "bandro"[tt] OR "bandro"[mh] OR "bandro"[rn] OR "bandro"[nm]) OR ("baxogar"[tiab] OR "baxogar"[tt] OR "baxogar"[mh] OR "baxogar"[rn] OR "baxogar"[nm]) OR ("bomanes"[tiab] OR "bomanes"[tt] OR "bomanes"[mh] OR "bomanes"[rn] OR "bomanes"[nm]) OR ("bonefrubit"[tiab] OR "bonefrubit"[tt] OR "bonefrubit"[mh] OR "bonefrubit"[rn] OR "bonefrubit"[nm]) OR ("bonefurbit"[tiab] OR "bonefurbit"[tt] OR "bonefurbit"[mh] OR "bonefurbit"[rn] OR "bonefurbit"[nm]) OR ("bonese"[tiab] OR "bonese"[tt] OR "bonese"[mh] OR "bonese"[rn] OR "bonese"[nm]) OR ("bonicid"[tiab] OR "bonicid"[tt] OR "bonicid"[mh] OR "bonicid"[rn] OR "bonicid"[nm]) OR ("bonmore"[tiab] OR "bonmore"[tt] OR "bonmore"[mh] OR "bonmore"[rn] OR "bonmore"[nm]) OR ("clastec"[tiab] OR "clastec"[tt] OR "clastec"[mh] OR "clastec"[rn] OR "clastec"[nm]) OR ("dronaval"[tiab] OR "dronaval"[tt] OR "dronaval"[mh] OR "dronaval"[rn] OR "dronaval"[nm]) OR ("fijical"[tiab] OR "fijical"[tt] OR "fijical"[mh] OR "fijical"[rn] OR "fijical"[nm]) OR ("holmevis"[tiab] OR "holmevis"[tt] OR "holmevis"[mh] OR "holmevis"[rn] OR "holmevis"[nm]) OR ("ibanat"[tiab] OR "ibanat"[tt] OR "ibanat"[mh] OR "ibanat"[rn] OR "ibanat"[nm]) OR ("ibandra"[tiab] OR "ibandra"[tt] OR "ibandra"[mh] OR "ibandra"[rn] OR "ibandra"[nm]) OR ("ibandrix"[tiab] OR "ibandrix"[tt] OR "ibandrix"[mh] OR "ibandrix"[rn] OR "ibandrix"[nm]) OR ("ibandronat"[tiab] OR "ibandronat"[tt] OR "ibandronat"[mh] OR "ibandronat"[rn] OR "ibandronat"[nm]) OR ("ibandronian"[tiab] OR "ibandronian"[tt] OR "ibandronian"[mh] OR "ibandronian"[rn] OR "ibandronian"[nm]) OR ("ibandronsav"[tiab] OR "ibandronsav"[tt] OR "ibandronsav"[mh] OR "ibandronsav"[rn] OR "ibandronsav"[nm]) OR ("ibanic"[tiab] OR "ibanic"[tt] OR "ibanic"[mh] OR "ibanic"[rn] OR "ibanic"[nm]) OR ("ibanos"[tiab] OR "ibanos"[tt] OR "ibanos"[mh] OR "ibanos"[rn] OR "ibanos"[nm]) OR ("ibone"[tiab] OR "ibone"[tt] OR "ibone"[mh] OR "ibone"[rn] OR "ibone"[nm]) OR ("ibrac"[tiab] OR "ibrac"[tt] OR "ibrac"[mh] OR "ibrac"[rn] OR "ibrac"[nm]) OR ("idena"[tiab] OR "idena"[tt] OR "idena"[mh] OR "idena"[rn] OR "idena"[nm]) OR ("ikametin"[tiab] OR "ikametin"[tt] OR "ikametin"[mh] OR "ikametin"[rn] OR "ikametin"[nm]) OR ("indrofar"[tiab] OR "indrofar"[tt] OR "indrofar"[mh] OR "indrofar"[rn] OR "indrofar"[nm]) OR ("ipexal"[tiab] OR "ipexal"[tt] OR "ipexal"[mh] OR "ipexal"[rn] OR "ipexal"[nm]) OR ("kefort"[tiab] OR "kefort"[tt] OR "kefort"[mh] OR "kefort"[rn] OR "kefort"[nm]) OR ("kemidat"[tiab] OR "kemidat"[tt] OR "kemidat"[mh] OR "kemidat"[rn] OR "kemidat"[nm]) OR ("licobondrat"[tiab] OR "licobondrat"[tt] OR "licobondrat"[mh] OR "licobondrat"[rn] OR "licobondrat"[nm]) OR ("meliba"[tiab] OR "meliba"[tt] OR "meliba"[mh] OR "meliba"[rn] OR "meliba"[nm]) OR ("nucodran"[tiab] OR "nucodran"[tt] OR "nucodran"[mh] OR "nucodran"[rn] OR "nucodran"[nm]) OR ("osagrand"[tiab] OR "osagrand"[tt] OR "osagrand"[mh] OR "osagrand"[rn] OR	1304

	"osagrand"[nm]) OR ("osbonelle"[tiab] OR "osbonelle"[tt] OR "osbonelle"[mh] OR "osbonelle"[rn] OR "osbonelle"[nm]) OR ("oseum"[tiab] OR "oseum"[tt] OR "oseum"[mh] OR "oseum"[rn] OR "oseum"[nm]) OR ("ossica"[tiab] OR "ossica"[tt] OR "ossica"[mh] OR "ossica"[rn] OR "ossica"[nm]) OR ("osteocalcit"[tiab] OR "osteocalcit"[tt] OR "osteocalcit"[mh] OR "osteocalcit"[rn] OR "osteocalcit"[nm]) OR ("osteolong"[tiab] OR "osteolong"[tt] OR "osteolong"[mh] OR "osteolong"[rn] OR "osteolong"[nm]) OR ("osteosyl"[tiab] OR "osteosyl"[tt] OR "osteosyl"[mh] OR "osteosyl"[rn] OR "osteosyl"[nm]) OR ("ostone"[tiab] OR "ostone"[tt] OR "ostone"[mh] OR "ostone"[rn] OR "ostone"[nm]) OR ("posclim"[tiab] OR "posclim"[tt] OR "posclim"[mh] OR "posclim"[rn] OR "posclim"[nm]) OR ("quodixor"[tiab] OR "quodixor"[tt] OR "quodixor"[mh] OR "quodixor"[rn] OR "quodixor"[nm]) OR ("recaxin"[tiab] OR "recaxin"[tt] OR "recaxin"[mh] OR "recaxin"[rn] OR "recaxin"[nm]) OR ("resormes"[tiab] OR "resormes"[tt] OR "resormes"[mh] OR "resormes"[rn] OR "resormes"[nm]) OR ("unomes"[tiab] OR "unomes"[tt] OR "unomes"[mh] OR "unomes"[rn] OR "unomes"[nm]) OR ("adromux"[tiab] OR "adromux"[tt] OR "adromux"[mh] OR "adromux"[rn] OR "adromux"[nm]) OR ("anabon"[tiab] OR "anabon"[tt] OR "anabon"[mh] OR "anabon"[rn] OR "anabon"[nm]) OR ("bandron"[tiab] OR "bandron"[tt] OR "bandron"[mh] OR "bandron"[rn] OR "bandron"[nm]) OR ("bantuc"[tiab] OR "bantuc"[tt] OR "bantuc"[mh] OR "bantuc"[rn] OR "bantuc"[nm]) OR ("baxogur"[tiab] OR "baxogur"[tt] OR "baxogur"[mh] OR "baxogur"[rn] OR "baxogur"[nm]) OR ("bonjenic"[tiab] OR "bonjenic"[tt] OR "bonjenic"[mh] OR "bonjenic"[rn] OR "bonjenic"[nm]) OR ("bonnedra"[tiab] OR "bonnedra"[tt] OR "bonnedra"[mh] OR "bonnedra"[rn] OR "bonnedra"[nm]) OR ("bonoste"[tiab] OR "bonoste"[tt] OR "bonoste"[mh] OR "bonoste"[rn] OR "bonoste"[nm]) OR ("darmas"[tiab] OR "darmas"[tt] OR "darmas"[mh] OR "darmas"[rn] OR "darmas"[nm]) OR ("disdual"[tiab] OR "disdual"[tt] OR "disdual"[mh] OR "disdual"[rn] OR "disdual"[nm]) OR ("elasterin"[tiab] OR "elasterin"[tt] OR "elasterin"[mh] OR "elasterin"[rn] OR "elasterin"[nm]) OR ("etanorden"[tiab] OR "etanorden"[tt] OR "etanorden"[mh] OR "etanorden"[rn] OR "etanorden"[nm]) OR ("femorel"[tiab] OR "femorel"[tt] OR "femorel"[mh] OR "femorel"[rn] OR "femorel"[nm]) OR ("haniban"[tiab] OR "haniban"[tt] OR "haniban"[mh] OR "haniban"[rn] OR "haniban"[nm]) OR ("ibagenit"[tiab] OR "ibagenit"[tt] OR "ibagenit"[mh] OR "ibagenit"[rn] OR "ibagenit"[nm]) OR ("ibames"[tiab] OR "ibames"[tt] OR "ibames"[mh] OR "ibames"[rn] OR "ibames"[nm]) OR ("ibamyl"[tiab] OR "ibamyl"[tt] OR "ibamyl"[mh] OR "ibamyl"[rn] OR "ibamyl"[nm]) OR ("Ibandroninezuur"[tiab] OR "Ibandroninezuur"[tt] OR "Ibandroninezuur"[mh] OR "Ibandroninezuur"[rn] OR "Ibandroninezuur"[nm]) OR ("Ibandronsav"[tiab] OR "Ibandronsav"[tt] OR "Ibandronsav"[mh] OR "Ibandronsav"[rn] OR "Ibandronsav"[nm]) OR ("ibandronsyre"[tiab] OR "ibandronsyre"[tt] OR "ibandronsyre"[mh] OR "ibandronsyre"[rn] OR "ibandronsyre"[nm]) OR ("ibanfos"[tiab] OR "ibanfos"[tt] OR "ibanfos"[mh] OR "ibanfos"[rn] OR "ibanfos"[nm]) OR ("ibanleg"[tiab] OR "ibanleg"[tt] OR "ibanleg"[mh] OR "ibanleg"[rn] OR "ibanleg"[nm]) OR ("ibannate"[tiab] OR "ibannate"[tt] OR "ibannate"[mh] OR "ibannate"[rn] OR "ibannate"[nm]) OR ("ibondro"[tiab] OR "ibondro"[tt] OR "ibondro"[mh] OR "ibondro"[rn] OR "ibondro"[nm]) OR ("ibostofar"[tiab] OR "ibostofar"[tt] OR "ibostofar"[mh] OR "ibostofar"[rn] OR "ibostofar"[nm]) OR ("idena"[tiab] OR "idena"[tt] OR "idena"[mh] OR "idena"[rn] OR "idena"[nm]) OR ("ikamentin"[tiab] OR "ikamentin"[tt] OR "ikamentin"[mh] OR "ikamentin"[rn] OR "ikamentin"[nm]) OR ("inostelid"[tiab] OR "inostelid"[tt] OR "inostelid"[mh] OR "inostelid"[rn] OR "inostelid"[nm]) OR ("kalosso"[tiab] OR "kalosso"[tt] OR "kalosso"[mh] OR "kalosso"[rn] OR "kalosso"[nm]) OR ("kefort"[tiab] OR "kefort"[tt] OR "kefort"[mh] OR "kefort"[rn] OR "kefort"[nm]) OR ("licobondrat"[tiab] OR "licobondrat"[tt] OR "licobondrat"[mh] OR "licobondrat"[rn] OR "licobondrat"[nm]) OR ("mirdezel"[tiab] OR "mirdezel"[tt] OR "mirdezel"[mh] OR "mirdezel"[rn] OR "mirdezel"[nm]) OR ("modifical"[tiab] OR "modifical"[tt] OR "modifical"[mh] OR "modifical"[rn] OR "modifical"[nm]) OR ("osma"[tiab] OR "osma"[tt] OR "osma"[mh] OR "osma"[rn] OR "osma"[nm]) OR	

	("osteonat"[tiab] OR "osteonat"[tt] OR "osteonat"[mh] OR "osteonat"[rn] OR "osteonat"[nm]) OR ("osteoviva"[tiab] OR "osteoviva"[tt] OR "osteoviva"[mh] OR "osteoviva"[rn] OR "osteoviva"[nm]) OR ("phacebonate"[tiab] OR "phacebonate"[tt] OR "phacebonate"[mh] OR "phacebonate"[rn] OR "phacebonate"[nm]) OR ("ratiban"[tiab] OR "ratiban"[tt] OR "ratiban"[mh] OR "ratiban"[rn] OR "ratiban"[nm]) OR ("recaxin"[tiab] OR "recaxin"[tt] OR "recaxin"[mh] OR "recaxin"[rn] OR "recaxin"[nm]) OR ("ribobandron"[tiab] OR "ribobandron"[tt] OR "ribobandron"[mh] OR "ribobandron"[rn] OR "ribobandron"[nm]) OR ("r-484"[tiab] OR "r-484"[tt] OR "r-484"[mh] OR "r-484"[rn] OR "r-484"[nm]) OR ("r484"[tiab] OR "r484"[tt] OR "r484"[mh] OR "r484"[rn] OR "r484"[nm]) OR ("114084-78-5"[tiab] OR "114084-78-5"[tt] OR "114084-78-5"[mh] OR "114084-78-5"[rn] OR "114084-78-5"[nm]) OR ("138844-81-2"[tiab] OR "138844-81-2"[tt] OR "138844-81-2"[mh] OR "138844-81-2"[rn] OR "138844-81-2"[nm]) OR ("138926-19-9"[tiab] OR "138926-19-9"[tt] OR "138926-19-9"[mh] OR "138926-19-9"[rn] OR "138926-19-9"[nm]))	
#12	(("zoledronic acid"[tiab] OR "zoledronic acid"[tt] OR "zoledronic acid"[mh] OR "zoledronic acid"[rn] OR "zoledronic acid"[nm]) OR ("zoledronate"[tiab] OR "zoledronate"[tt] OR "zoledronate"[mh] OR "zoledronate"[rn] OR "zoledronate"[nm]) OR ("Aclasta"[tiab] OR "Aclasta"[tt] OR "Aclasta"[mh] OR "Aclasta"[rn] OR "Aclasta"[nm]) OR ("Reclast"[tiab] OR "Reclast"[tt] OR "Reclast"[mh] OR "Reclast"[rn] OR "Reclast"[nm]) OR ("cgp-42446"[tiab] OR "cgp-42446"[tt] OR "cgp-42446"[mh] OR "cgp-42446"[rn] OR "cgp-42446"[nm]) OR ("cgp42446"[tiab] OR "cgp42446"[tt] OR "cgp42446"[mh] OR "cgp42446"[rn] OR "cgp42446"[nm]) OR ("cgp-42446a"[tiab] OR "cgp-42446a"[tt] OR "cgp-42446a"[mh] OR "cgp-42446a"[rn] OR "cgp-42446a"[nm]) OR ("cgp42446a"[tiab] OR "cgp42446a"[tt] OR "cgp42446a"[mh] OR "cgp42446a"[rn] OR "cgp42446a"[nm]) OR ("orazol"[tiab] OR "orazol"[tt] OR "orazol"[mh] OR "orazol"[rn] OR "orazol"[nm]) OR ("zol-446"[tiab] OR "zol-446"[tt] OR "zol-446"[mh] OR "zol-446"[rn] OR "zol-446"[nm]) OR ("zol446"[tiab] OR "zol446"[tt] OR "zol446"[mh] OR "zol446"[rn] OR "zol446"[nm]) OR ("zomera"[tiab] OR "zomera"[tt] OR "zomera"[mh] OR "zomera"[rn] OR "zomera"[nm]) OR ("zometa"[tiab] OR "zometa"[tt] OR "zometa"[mh] OR "zometa"[rn] OR "zometa"[nm]) OR ("blaztere"[tiab] OR "blaztere"[tt] OR "blaztere"[mh] OR "blaztere"[rn] OR "blaztere"[nm]) OR ("bolenic"[tiab] OR "bolenic"[tt] OR "bolenic"[mh] OR "bolenic"[rn] OR "bolenic"[nm]) OR ("boncur"[tiab] OR "boncur"[tt] OR "boncur"[mh] OR "boncur"[rn] OR "boncur"[nm]) OR ("celdron"[tiab] OR "celdron"[tt] OR "celdron"[mh] OR "celdron"[rn] OR "celdron"[nm]) OR ("desibon"[tiab] OR "desibon"[tt] OR "desibon"[mh] OR "desibon"[rn] OR "desibon"[nm]) OR ("drometa"[tiab] OR "drometa"[tt] OR "drometa"[mh] OR "drometa"[rn] OR "drometa"[nm]) OR ("eriophos"[tiab] OR "eriophos"[tt] OR "eriophos"[mh] OR "eriophos"[rn] OR "eriophos"[nm]) OR ("fayton"[tiab] OR "fayton"[tt] OR "fayton"[mh] OR "fayton"[rn] OR "fayton"[nm]) OR ("kaliksir"[tiab] OR "kaliksir"[tt] OR "kaliksir"[mh] OR "kaliksir"[rn] OR "kaliksir"[nm]) OR ("ledron"[tiab] OR "ledron"[tt] OR "ledron"[mh] OR "ledron"[rn] OR "ledron"[nm]) OR ("osporil"[tiab] OR "osporil"[tt] OR "osporil"[mh] OR "osporil"[rn] OR "osporil"[nm]) OR ("ostezolen"[tiab] OR "ostezolen"[tt] OR "ostezolen"[mh] OR "ostezolen"[rn] OR "ostezolen"[nm]) OR ("rionit"[tiab] OR "rionit"[tt] OR "rionit"[mh] OR "rionit"[rn] OR "rionit"[nm]) OR ("simpla"[tiab] OR "simpla"[tt] OR "simpla"[mh] OR "simpla"[rn] OR "simpla"[nm]) OR ("sinresor"[tiab] OR "sinresor"[tt] OR "sinresor"[mh] OR "sinresor"[rn] OR "sinresor"[nm]) OR ("steozol"[tiab] OR "steozol"[tt] OR "steozol"[mh] OR "steozol"[rn] OR "steozol"[nm]) OR ("synblasta"[tiab] OR "synblasta"[tt] OR "synblasta"[mh] OR "synblasta"[rn] OR "synblasta"[nm]) OR ("syndronic"[tiab] OR "syndronic"[tt] OR "syndronic"[mh] OR "syndronic"[rn] OR "syndronic"[nm]) OR ("varidronico"[tiab] OR "varidronico"[tt] OR "varidronico"[mh] OR "varidronico"[rn] OR "varidronico"[nm]) OR ("zelinda"[tiab] OR "zelinda"[tt] OR "zelinda"[mh] OR "zelinda"[rn] OR "zelinda"[nm]) OR ("zidolamin"[tiab] OR	6385

	"zidolamin"[tt] OR "zidolamin"[mh] OR "zidolamin"[rn] OR "zidolamin"[nm]) OR ("zidronic"[tiab] OR "zidronic"[tt] OR "zidronic"[mh] OR "zidronic"[rn] OR "zidronic"[nm]) OR ("ziduvin"[tiab] OR "ziduvin"[tt] OR "ziduvin"[mh] OR "ziduvin"[rn] OR "ziduvin"[nm]) OR ("zinvel"[tiab] OR "zinvel"[tt] OR "zinvel"[mh] OR "zinvel"[rn] OR "zinvel"[nm]) OR ("zobone"[tiab] OR "zobone"[tt] OR "zobone"[mh] OR "zobone"[rn] OR "zobone"[nm]) OR ("zobonic"[tiab] OR "zobonic"[tt] OR "zobonic"[mh] OR "zobonic"[rn] OR "zobonic"[nm]) OR ("zolacitor"[tiab] OR "zolacitor"[tt] OR "zolacitor"[mh] OR "zolacitor"[rn] OR "zolacitor"[nm]) OR ("zolako"[tiab] OR "zolako"[tt] OR "zolako"[mh] OR "zolako"[rn] OR "zolako"[nm]) OR ("zoledro"[tiab] OR "zoledro"[tt] OR "zoledro"[mh] OR "zoledro"[rn] OR "zoledro"[nm]) OR ("zoledreenos"[tiab] OR "zoledreenos"[tt] OR "zoledreenos"[mh] OR "zoledreenos"[rn] OR "zoledreenos"[nm]) OR ("zoledrin"[tiab] OR "zoledrin"[tt] OR "zoledrin"[mh] OR "zoledrin"[rn] OR "zoledrin"[nm]) OR ("zoledronate"[tiab] OR "zoledronate"[tt] OR "zoledronate"[mh] OR "zoledronate"[rn] OR "zoledronate"[nm]) OR ("zoledronsyre"[tiab] OR "zoledronsyre"[tt] OR "zoledronsyre"[mh] OR "zoledronsyre"[rn] OR "zoledronsyre"[nm]) OR ("zolenat"[tiab] OR "zolenat"[tt] OR "zolenat"[mh] OR "zolenat"[rn] OR "zolenat"[nm]) OR ("zolenic"[tiab] OR "zolenic"[tt] OR "zolenic"[mh] OR "zolenic"[rn] OR "zolenic"[nm]) OR ("zoletalis"[tiab] OR "zoletalis"[tt] OR "zoletalis"[mh] OR "zoletalis"[rn] OR "zoletalis"[nm]) OR ("zoletech"[tiab] OR "zoletech"[tt] OR "zoletech"[mh] OR "zoletech"[rn] OR "zoletech"[nm]) OR ("zolira"[tiab] OR "zolira"[tt] OR "zolira"[mh] OR "zolira"[rn] OR "zolira"[nm]) OR ("zomebon"[tiab] OR "zomebon"[tt] OR "zomebon"[mh] OR "zomebon"[rn] OR "zomebon"[nm]) OR ("zomedron"[tiab] OR "zomedron"[tt] OR "zomedron"[mh] OR "zomedron"[rn] OR "zomedron"[nm]) OR ("zomera"[tiab] OR "zomera"[tt] OR "zomera"[mh] OR "zomera"[rn] OR "zomera"[nm]) OR ("zometa"[tiab] OR "zometa"[tt] OR "zometa"[mh] OR "zometa"[rn] OR "zometa"[nm]) OR ("zomikos"[tiab] OR "zomikos"[tt] OR "zomikos"[mh] OR "zomikos"[rn] OR "zomikos"[nm]) OR ("zuorui"[tiab] OR "zuorui"[tt] OR "zuorui"[mh] OR "zuorui"[rn] OR "zuorui"[nm]) OR ("zyolix"[tiab] OR "zyolix"[tt] OR "zyolix"[mh] OR "zyolix"[rn] OR "zyolix"[nm]) OR ("cenozoic"[tiab] OR "cenozoic"[tt] OR "cenozoic"[mh] OR "cenozoic"[rn] OR "cenozoic"[nm]) OR ("desinobon"[tiab] OR "desinobon"[tt] OR "desinobon"[mh] OR "desinobon"[rn] OR "desinobon"[nm]) OR ("indaferil"[tiab] OR "indaferil"[tt] OR "indaferil"[mh] OR "indaferil"[rn] OR "indaferil"[nm]) OR ("midronic"[tiab] OR "midronic"[tt] OR "midronic"[mh] OR "midronic"[rn] OR "midronic"[nm]) OR ("leuzotev"[tiab] OR "leuzotev"[tt] OR "leuzotev"[mh] OR "leuzotev"[rn] OR "leuzotev"[nm]) OR ("tevadronic"[tiab] OR "tevadronic"[tt] OR "tevadronic"[mh] OR "tevadronic"[rn] OR "tevadronic"[nm]) OR ("zacindate"[tiab] OR "zacindate"[tt] OR "zacindate"[mh] OR "zacindate"[rn] OR "zacindate"[nm]) OR ("zalit"[tiab] OR "zalit"[tt] OR "zalit"[mh] OR "zalit"[rn] OR "zalit"[nm]) OR ("zofaden"[tiab] OR "zofaden"[tt] OR "zofaden"[mh] OR "zofaden"[rn] OR "zofaden"[nm]) OR ("zolacin"[tiab] OR "zolacin"[tt] OR "zolacin"[mh] OR "zolacin"[rn] OR "zolacin"[nm]) OR ("zoldria"[tiab] OR "zoldria"[tt] OR "zoldria"[mh] OR "zoldria"[rn] OR "zoldria"[nm]) OR ("zoledo"[tiab] OR "zoledo"[tt] OR "zoledo"[mh] OR "zoledo"[rn] OR "zoledo"[nm]) OR ("zolecan"[tiab] OR "zolecan"[tt] OR "zolecan"[mh] OR "zolecan"[rn] OR "zolecan"[nm]) OR ("zoledronsav"[tiab] OR "zoledronsav"[tt] OR "zoledronsav"[mh] OR "zoledronsav"[rn] OR "zoledronsav"[nm]) OR ("zolenia"[tiab] OR "zolenia"[tt] OR "zolenia"[mh] OR "zolenia"[rn] OR "zolenia"[nm]) OR ("zortila"[tiab] OR "zortila"[tt] OR "zortila"[mh] OR "zortila"[rn] OR "zortila"[nm]) OR ("118072-93-8"[tiab] OR "118072-93-8"[tt] OR "118072-93-8"[mh] OR "118072-93-8"[rn] OR "118072-93-8"[nm]) OR ("131654-46-1"[tiab] OR "131654-46-1"[tt] OR "131654-46-1"[mh] OR "131654-46-1"[rn] OR "131654-46-1"[nm]) OR ("165800-06-6"[tiab] OR "165800-06-6"[tt] OR "165800-06-6"[mh] OR "165800-06-6"[rn] OR "165800-06-6"[nm]) OR ("165800-07-7"[tiab] OR "165800-07-7"[tt] OR "165800-07-7"[mh] OR "165800-07-7"[rn] OR "165800-07-7"[nm]))	
#13	"Denosumab"[mesh:noexp]	1652

#14	(("denosumab"[tiab] OR "denosumab"[tt] OR "denosumab"[mh] OR "denosumab"[rn] OR "denosumab"[nm]) OR ("amg 162"[tiab] OR "amg 162"[tt] OR "amg 162"[mh] OR "amg 162"[rn] OR "amg 162"[nm]) OR ("amg162"[tiab] OR "amg162"[tt] OR "amg162"[mh] OR "amg162"[rn] OR "amg162"[nm]) OR ("amgiva"[tiab] OR "amgiva"[tt] OR "amgiva"[mh] OR "amgiva"[rn] OR "amgiva"[nm]) OR ("prolia"[tiab] OR "prolia"[tt] OR "prolia"[mh] OR "prolia"[rn] OR "prolia"[nm]) OR ("615258-40-7"[tiab] OR "615258-40-7"[tt] OR "615258-40-7"[mh] OR "615258-40-7"[rn] OR "615258-40-7"[nm]))	3095
#15	#13 OR #14	3095
#16	"Raloxifene Hydrochloride"[mesh:noexp]	2615
#17	(("Raloxifene"[tiab] OR "Raloxifene"[tt] OR "Raloxifene"[mh] OR "Raloxifene"[rn] OR "Raloxifene"[nm]) OR ("LY139481"[tiab] OR "LY139481"[tt] OR "LY139481"[mh] OR "LY139481"[rn] OR "LY139481"[nm]) OR ("LY-139481"[tiab] OR "LY-139481"[tt] OR "LY-139481"[mh] OR "LY-139481"[rn] OR "LY-139481"[nm]) OR ("bonmax"[tiab] OR "bonmax"[tt] OR "bonmax"[mh] OR "bonmax"[rn] OR "bonmax"[nm]) OR ("celvista"[tiab] OR "celvista"[tt] OR "celvista"[mh] OR "celvista"[rn] OR "celvista"[nm]) OR ("evista"[tiab] OR "evista"[tt] OR "evista"[mh] OR "evista"[rn] OR "evista"[nm]) OR ("keoxifene"[tiab] OR "keoxifene"[tt] OR "keoxifene"[mh] OR "keoxifene"[rn] OR "keoxifene"[nm]) OR ("loxar"[tiab] OR "loxar"[tt] OR "loxar"[mh] OR "loxar"[rn] OR "loxar"[nm]) OR ("loxifen"[tiab] OR "loxifen"[tt] OR "loxifen"[mh] OR "loxifen"[rn] OR "loxifen"[nm]) OR ("ly-156758"[tiab] OR "ly-156758"[tt] OR "ly-156758"[mh] OR "ly-156758"[rn] OR "ly-156758"[nm]) OR ("ly156758"[tiab] OR "ly156758"[tt] OR "ly156758"[mh] OR "ly156758"[rn] OR "ly156758"[nm]) OR ("ly139481"[tiab] OR "ly139481"[tt] OR "ly139481"[mh] OR "ly139481"[rn] OR "ly139481"[nm]) OR ("ly-139481"[tiab] OR "ly-139481"[tt] OR "ly-139481"[mh] OR "ly-139481"[rn] OR "ly-139481"[nm]) OR ("raxeto"[tiab] OR "raxeto"[tt] OR "raxeto"[mh] OR "raxeto"[rn] OR "raxeto"[nm]) OR ("evista"[tiab] OR "evista"[tt] OR "evista"[mh] OR "evista"[rn] OR "evista"[nm]) OR ("fluken"[tiab] OR "fluken"[tt] OR "fluken"[mh] OR "fluken"[rn] OR "fluken"[nm]) OR ("gynista"[tiab] OR "gynista"[tt] OR "gynista"[mh] OR "gynista"[rn] OR "gynista"[nm]) OR ("osteoclax"[tiab] OR "osteoclax"[tt] OR "osteoclax"[mh] OR "osteoclax"[rn] OR "osteoclax"[nm]) OR ("osteya"[tiab] OR "osteya"[tt] OR "osteya"[mh] OR "osteya"[rn] OR "osteya"[nm]) OR ("ostiral"[tiab] OR "ostiral"[tt] OR "ostiral"[mh] OR "ostiral"[rn] OR "ostiral"[nm]) OR ("ralosto"[tiab] OR "ralosto"[tt] OR "ralosto"[mh] OR "ralosto"[rn] OR "ralosto"[nm]) OR ("raloxa"[tiab] OR "raloxa"[tt] OR "raloxa"[mh] OR "raloxa"[rn] OR "raloxa"[nm]) OR ("ronixifeno"[tiab] OR "ronixifeno"[tt] OR "ronixifeno"[mh] OR "ronixifeno"[rn] OR "ronixifeno"[nm]) OR ("aloxif"[tiab] OR "aloxif"[tt] OR "aloxif"[mh] OR "aloxif"[rn] OR "aloxif"[nm]) OR ("optruma"[tiab] OR "optruma"[tt] OR "optruma"[mh] OR "optruma"[rn] OR "optruma"[nm]) OR ("oxilar"[tiab] OR "oxilar"[tt] OR "oxilar"[mh] OR "oxilar"[rn] OR "oxilar"[nm]) OR ("raloksifen"[tiab] OR "raloksifen"[tt] OR "raloksifen"[mh] OR "raloksifen"[rn] OR "raloksifen"[nm]) OR ("ralomeer"[tiab] OR "ralomeer"[tt] OR "ralomeer"[mh] OR "ralomeer"[rn] OR "ralomeer"[nm]) OR ("ralopharm"[tiab] OR "ralopharm"[tt] OR "ralopharm"[mh] OR "ralopharm"[rn] OR "ralopharm"[nm]) OR ("ralover"[tiab] OR "ralover"[tt] OR "ralover"[mh] OR "ralover"[rn] OR "ralover"[nm]) OR ("ralox"[tiab] OR "ralox"[tt] OR "ralox"[mh] OR "ralox"[rn] OR "ralox"[nm]) OR ("raloxa"[tiab] OR "raloxa"[tt] OR "raloxa"[mh] OR "raloxa"[rn] OR "raloxa"[nm]) OR ("raloxibone"[tiab] OR "raloxibone"[tt] OR "raloxibone"[mh] OR "raloxibone"[rn] OR "raloxibone"[nm]) OR ("raloxiep"[tiab] OR "raloxiep"[tt] OR "raloxiep"[mh] OR "raloxiep"[rn] OR "raloxiep"[nm]) OR ("raloxifen"[tiab] OR "raloxifen"[tt] OR "raloxifen"[mh] OR "raloxifen"[rn] OR "raloxifen"[nm]) OR ("raloxstar"[tiab] OR "raloxstar"[tt] OR "raloxstar"[mh] OR "raloxstar"[rn] OR "raloxstar"[nm]) OR ("raloxten"[tiab] OR "raloxten"[tt] OR "raloxten"[mh] OR "raloxten"[rn] OR "raloxten"[nm]) OR ("82640-04-8"[tiab] OR "82640-04-8"[tt] OR "82640-04-8"[mh] OR "82640-04-	3981

	8"[rn] OR "82640-04-8"[nm]) OR ("84449-90-1"[tiab] OR "84449-90-1"[tt] OR "84449-90-1"[mh] OR "84449-90-1"[rn] OR "84449-90-1"[nm]))	
#18	#16 OR #17	3981
#19	(("Abaloparatide"[tiab] OR "Abaloparatide"[tt] OR "Abaloparatide"[mh] OR "Abaloparatide"[rn] OR "Abaloparatide"[nm]) OR ("BA058"[tiab] OR "BA058"[tt] OR "BA058"[mh] OR "BA058"[rn] OR "BA058"[nm]) OR ("BA-058"[tiab] OR "BA-058"[tt] OR "BA-058"[mh] OR "BA-058"[rn] OR "BA-058"[nm]) OR ("bim-44058"[tiab] OR "bim-44058"[tt] OR "bim-44058"[mh] OR "bim-44058"[rn] OR "bim-44058"[nm]) OR ("bim44058"[tiab] OR "bim44058"[tt] OR "bim44058"[mh] OR "bim44058"[rn] OR "bim44058"[nm]) OR ("247062-33-5"[tiab] OR "247062-33-5"[tt] OR "247062-33-5"[mh] OR "247062-33-5"[rn] OR "247062-33-5"[nm]))	148
#20	"Cathepsin K/Antagonists and Inhibitors"[mesh:noexp]	226
#21	(("Odanacatib"[tiab] OR "Odanacatib"[tt] OR "Odanacatib"[mh] OR "Odanacatib"[rn] OR "Odanacatib"[nm]) OR ("MK0822"[tiab] OR "MK0822"[tt] OR "MK0822"[mh] OR "MK0822"[rn] OR "MK0822"[nm]) OR ("MK-0822"[tiab] OR "MK-0822"[tt] OR "MK-0822"[mh] OR "MK-0822"[rn] OR "MK-0822"[nm]) OR ("mk822"[tiab] OR "mk822"[tt] OR "mk822"[mh] OR "mk822"[rn] OR "mk822"[nm]) OR ("mk-822"[tiab] OR "mk-822"[tt] OR "mk-822"[mh] OR "mk-822"[rn] OR "mk-822"[nm]) OR ("603139-19-1"[tiab] OR "603139-19-1"[tt] OR "603139-19-1"[mh] OR "603139-19-1"[rn] OR "603139-19-1"[nm]))	200
#22	(Cathepsin K inhibitor*[tiab] OR Cathepsin K inhibitor*[tt])	312
#23	#20 OR #21 OR #22	504
#24	“terrosa or RGB-10 or RGB10 or movymia”	18
#25	#1 OR #4 OR #7 OR #10 OR #11 OR #12 OR #15 OR #18 OR #19 OR #23 OR #24	22607
#26	"randomized controlled trial"[pt:noexp] OR "randomized controlled trials as topic"[mesh:noexp]	639539
#27	"controlled clinical trial"[pt:noexp]	93778
#28	(random*[ti] OR random*[tt])	228713
#29	"placebo"[tiab]	215135
#30	"DRUG THERAPY"[sh:noexp]	2222854
#31	random*[tiab]	1145795
#32	"trial"[tiab]	604807
#33	"groups"[tiab]	2098044
#34	#26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33	5062308

#35	("animals"[MeSH Terms:noexp] OR "animals"[All Fields]) NOT (("animals"[MeSH Terms:noexp] OR "animals"[All Fields]) AND ((((((((((((((("human s"[All Fields] OR "humane"[All Fields]) OR "humanely"[All Fields]) OR "humaneness"[All Fields]) OR "humanism"[MeSH Terms]) OR "humanism"[All Fields]) OR "humanities"[MeSH Terms]) OR "humanities"[All Fields]) OR "humanity"[All Fields]) OR "humanity s"[All Fields]) OR "humanization"[All Fields]) OR "humanize"[All Fields]) OR "humanizes"[All Fields]) OR "humanizing"[All Fields]) OR "humanness"[All Fields]) OR "humans"[MeSH Terms]) OR "humans"[All Fields]) OR "human"[All Fields]))	4420784
#36	#34 NOT #35	4444815
#37	"Osteoporosis"[mesh]	55690
#38	"Bone Diseases, Metabolic"[mesh:noexp]	7813
#39	"Bone Density"[mesh:noexp]	53207
#40	"Fractures, Bone"[mesh]	183874
#41	((osteoporo*[tiab] OR osteoporo*[tt] OR osteoporo*[mh]) OR (osteo-poro*[tiab] OR osteo-poro*[tt] OR osteo-poro*[mh]))	93044
#42	((fragil*[tiab] OR fragil*[tt]) AND ((fractur*[tiab] OR fractur*[tt]) OR (break*[tiab] OR break*[tt])))	7876
#43	((osteoporotic decalcif*[tiab] OR osteoporotic decalcif*[tt]) OR (patholog* decalcif$[tiab] OR patholog* decalcif$[tt]) OR (osteopeni*[tiab] OR osteopeni*[tt]))	11093
#44	((bone mineral dens*[tiab] OR bone mineral dens*[tt] OR bone mineral dens*[mh]) OR ("bone loss"[tiab] OR "bone loss"[tt] OR "bone loss"[mh]) OR (bone fragil*[tiab] OR bone fragil*[tt] OR bone fragil*[mh]))	67665
#45	("BMD"[tiab] OR "BMD"[tt])	29318
#46	(fractur*[tiab] OR fractur*[tt])	262477
#47	((bone*[tiab] OR bone*[tt] OR bone*[mh]) AND (("density"[tiab] OR "density"[tt] OR "density"[mh]) OR (break*[tiab] OR break*[tt] OR break*[mh]) OR ("porosity"[tiab] OR "porosity"[tt] OR "porosity"[mh]) OR ("porotic"[tiab] OR "porotic"[tt] OR "porotic"[mh]) OR (decalcif*[tiab] OR decalcif*[tt] OR decalcif*[mh])))	98532
#48	#37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47	447804
#49	#25 AND #36 AND #48	7715
#50	**#49 Filters: from 2018 - 2020**	**911**

Table 65: Summary of EMBASE searches

No.	Embase Search	Search hits
#1	'romosozumab'/exp OR romosozumab OR romosozumab:ti,ab,de,rn,tn OR evenity:ti,ab,de,rn,tn OR amg785:ti,ab,de,rn,tn OR 'amg 785':ti,ab,de,rn,tn OR 'cdp 7851':ti,ab,de,rn,tn OR cdp7851:ti,ab,de,rn,tn OR '909395 70 6':ti,ab,de,rn,tn	595
#2	parathyroid hormone[1-34]'/exp OR teriparatide:ti,ab,de,rn,tn OR forteo:ti,ab,de,rn,tn OR forsteo:ti,ab,de,rn,tn OR 'chs 13340':ti,ab,de,rn,tn OR chs13340:ti,ab,de,rn,tn OR 'ly 333334':ti,ab,de,rn,tn OR ly333334:ti,ab,de,rn,tn OR parathar:ti,ab,de,rn,tn OR 'parathormone 1 34':ti,ab,de,rn,tn OR 'parathyroid hormone 1-34':ti,ab,de,rn,tn OR 'pth[1-34]':ti,ab,de,rn,tn OR 'sun-e3001':ti,ab,de,rn,tn OR 'sune3001':ti,ab,de,rn,tn OR '12583 68 5':ti,ab,de,rn,tn OR '52232 67 4':ti,ab,de,rn,tn	8695
#3	alendronic AND 'acid'/exp OR (alendronic:ti,ab,de,rn,tn AND acid:ti,ab,de,rn,tn) OR alendronate:ti,ab,de,rn,tn OR alenato:ti,ab,de,rn,tn OR alend:ti,ab,de,rn,tn OR alendros:ti,ab,de,rn,tn OR alovell:ti,ab,de,rn,tn OR arendal:ti,ab,de,rn,tn OR bifemelan:ti,ab,de,rn,tn OR bifosa:ti,ab,de,rn,tn OR bonapex:ti,ab,de,rn,tn OR defixal:ti,ab,de,rn,tn OR dronal:ti,ab,de,rn,tn OR endronax:ti,ab,de,rn,tn OR eucalen:ti,ab,de,rn,tn OR fixopan:ti,ab,de,rn,tn OR fosalan:ti,ab,de,rn,tn OR fosamax:ti,ab,de,rn,tn OR fosmin:ti,ab,de,rn,tn OR fosval:ti,ab,de,rn,tn OR marvil:ti,ab,de,rn,tn OR maxibone:ti,ab,de,rn,tn OR 'mk-0217':ti,ab,de,rn,tn OR 'mk 217':ti,ab,de,rn,tn OR mk0217:ti,ab,de,rn,tn OR mk217:ti,ab,de,rn,tn OR neobon:ti,ab,de,rn,tn OR oncalst:ti,ab,de,rn,tn OR onclast:ti,ab,de,rn,tn OR osdron:ti,ab,de,rn,tn OR osdronat:ti,ab,de,rn,tn OR oseotenk:ti,ab,de,rn,tn OR osficar:ti,ab,de,rn,tn OR oslene:ti,ab,de,rn,tn OR osteofar:ti,ab,de,rn,tn OR osteofos:ti,ab,de,rn,tn OR osteopor:ti,ab,de,rn,tn OR osteosan:ti,ab,de,rn,tn OR osteovan:ti,ab,de,rn,tn OR osticalcin:ti,ab,de,rn,tn OR porosal:ti,ab,de,rn,tn OR teiroc:ti,ab,de,rn,tn OR tibolene:ti,ab,de,rn,tn OR voroste:ti,ab,de,rn,tn OR fosavance:ti,ab,de,rn,tn OR adrovance:ti,ab,de,rn,tn OR vantavo:ti,ab,de,rn,tn OR binosto:ti,ab,de,rn,tn OR mylan:ti,ab,de,rn,tn OR adronat:ti,ab,de,rn,tn OR alendro:ti,ab,de,rn,tn OR alendraccord:ti,ab,de,rn,tn OR alendrobell:ti,ab,de,rn,tn OR 'alendrocor 10':ti,ab,de,rn,tn OR 'densate 70':ti,ab,de,rn,tn OR 'dronalen plus':ti,ab,de,rn,tn OR ossmax:ti,ab,de,rn,tn OR '66376 36 1':ti,ab,de,rn,tn	17494

#4	risedronic AND 'acid'/exp OR (risedronic:ti,ab,de,rn,tn AND acid:ti,ab,de,rn,tn) OR actonel:ti,ab,de,rn,tn OR atelvia:ti,ab,de,rn,tn OR 'ne 58095':ti,ab,de,rn,tn OR ne58095:ti,ab,de,rn,tn OR optinate:ti,ab,de,rn,tn OR risedronate:ti,ab,de,rn,tn OR acris:ti,ab,de,rn,tn OR risedro:ti,ab,de,rn,tn OR benet:ti,ab,de,rn,tn OR 'co risedrocal combo kit':ti,ab,de,rn,tn OR aktonate:ti,ab,de,rn,tn OR bonna:ti,ab,de,rn,tn OR cladronate:ti,ab,de,rn,tn OR ductonar:ti,ab,de,rn,tn OR goyart:ti,ab,de,rn,tn OR melenor:ti,ab,de,rn,tn OR ostenel:ti,ab,de,rn,tn OR osteodronate:ti,ab,de,rn,tn OR 'ribastamin duo rigat':ti,ab,de,rn,tn OR risate:ti,ab,de,rn,tn OR risedron:ti,ab,de,rn,tn OR risedrogen:ti,ab,de,rn,tn OR risendronat:ti,ab,de,rn,tn OR risemylan:ti,ab,de,rn,tn OR risendal:ti,ab,de,rn,tn OR isendros:ti,ab,de,rn,tn OR risetab:ti,ab,de,rn,tn OR salost:ti,ab,de,rn,tn OR tracost:ti,ab,de,rn,tn OR acrel:ti,ab,de,rn,tn OR actomax:ti,ab,de,rn,tn OR actojenic:ti,ab,de,rn,tn OR actokit:ti,ab,de,rn,tn OR arilex:ti,ab,de,rn,tn OR atconate:ti,ab,de,rn,tn OR bondapen:ti,ab,de,rn,tn OR boneact:ti,ab,de,rn,tn OR boncur:ti,ab,de,rn,tn OR bonmate:ti,ab,de,rn,tn OR bontonel:ti,ab,de,rn,tn OR bontrol:ti,ab,de,rn,tn OR claronate:ti,ab,de,rn,tn OR enospag:ti,ab,de,rn,tn OR fodren:ti,ab,de,rn,tn OR juverital:ti,ab,de,rn,tn OR medeoros:ti,ab,de,rn,tn OR miosen:ti,ab,de,rn,tn OR natalox:ti,ab,de,rn,tn OR norifax:ti,ab,de,rn,tn OR norsed:ti,ab,de,rn,tn OR osodens:ti,ab,de,rn,tn OR osteoron:ti,ab,de,rn,tn OR ostron:ti,ab,de,rn,tn OR pexalit:ti,ab,de,rn,tn OR tentop:ti,ab,de,rn,tn OR resorpate:ti,ab,de,rn,tn OR retonel:ti,ab,de,rn,tn OR ribastamin:ti,ab,de,rn,tn OR ribidron:ti,ab,de,rn,tn OR ribone:ti,ab,de,rn,tn OR richbone:ti,ab,de,rn,tn OR ridbone:ti,ab,de,rn,tn OR ridron:ti,ab,de,rn,tn OR ridrone:ti,ab,de,rn,tn OR risadican:ti,ab,de,rn,tn OR risbon:ti,ab,de,rn,tn OR risebon:ti,ab,de,rn,tn OR risebone:ti,ab,de,rn,tn OR risedon:ti,ab,de,rn,tn OR risedreenos:ti,ab,de,rn,tn OR risedronaat:ti,ab,de,rn,tn OR riselib:ti,ab,de,rn,tn OR risemed:ti,ab,de,rn,tn OR risedrenos:ti,ab,de,rn,tn OR risenex:ti,ab,de,rn,tn OR risenil:ti,ab,de,rn,tn OR riseto:ti,ab,de,rn,tn OR risetron:ti,ab,de,rn,tn OR resmyl:ti,ab,de,rn,tn OR risofos:ti,ab,de,rn,tn OR risonate:ti,ab,de,rn,tn OR risonato:ti,ab,de,rn,tn OR risostad:ti,ab,de,rn,tn OR ristonat:ti,ab,de,rn,tn OR sedron:ti,ab,de,rn,tn OR seralis:ti,ab,de,rn,tn OR tecnodron:ti,ab,de,rn,tn OR tevanel:ti,ab,de,rn,tn OR varibona:ti,ab,de,rn,tn OR norifaz:ti,ab,de,rn,tn OR zectoel:ti,ab,de,rn,tn OR acridon:ti,ab,de,rn,tn OR ridroqueen:ti,ab,de,rn,tn OR '105462 24 6':ti,ab,de,rn,tn OR '122458 82 6':ti,ab,de,rn,tn	8216

#5	ibandronic AND 'acid'/exp OR ((ibandronate:ti,ab,de,rn,tn OR ibandronic:ti,ab,de,rn,tn) AND acid:ti,ab,de,rn,tn) OR bonviva:ti,ab,de,rn,tn OR bondronat:ti,ab,de,rn,tn OR bondronate:ti,ab,de,rn,tn OR boniva:ti,ab,de,rn,tn OR destara:ti,ab,de,rn,tn OR 'bm 210955':ti,ab,de,rn,tn OR bm210955:ti,ab,de,rn,tn OR bondenza:ti,ab,de,rn,tn OR iasibon:ti,ab,de,rn,tn OR ibandronico:ti,ab,de,rn,tn OR alvodron:ti,ab,de,rn,tn OR alvodronic:ti,ab,de,rn,tn OR bandro:ti,ab,de,rn,tn OR baxogar:ti,ab,de,rn,tn OR bomanes:ti,ab,de,rn,tn OR bonefrubit:ti,ab,de,rn,tn OR bonefurbit:ti,ab,de,rn,tn OR bonese:ti,ab,de,rn,tn OR bonicid:ti,ab,de,rn,tn OR bonmore:ti,ab,de,rn,tn OR clastec:ti,ab,de,rn,tn OR dronaval:ti,ab,de,rn,tn OR fijical:ti,ab,de,rn,tn OR holmevis:ti,ab,de,rn,tn OR ibanat:ti,ab,de,rn,tn OR ibandra:ti,ab,de,rn,tn OR ibandrix:ti,ab,de,rn,tn OR ibandronat:ti,ab,de,rn,tn OR ibandronian:ti,ab,de,rn,tn OR ibanic:ti,ab,de,rn,tn OR ibanos:ti,ab,de,rn,tn OR ibone:ti,ab,de,rn,tn OR ibrac:ti,ab,de,rn,tn OR ikametin:ti,ab,de,rn,tn OR indrofar:ti,ab,de,rn,tn OR ipexal:ti,ab,de,rn,tn OR kemidat:ti,ab,de,rn,tn OR meliba:ti,ab,de,rn,tn OR nucodran:ti,ab,de,rn,tn OR osagrand:ti,ab,de,rn,tn OR osbonelle:ti,ab,de,rn,tn OR oseum:ti,ab,de,rn,tn OR ossica:ti,ab,de,rn,tn OR osteocalcit:ti,ab,de,rn,tn OR osteolong:ti,ab,de,rn,tn OR osteosyl:ti,ab,de,rn,tn OR ostone:ti,ab,de,rn,tn OR posclim:ti,ab,de,rn,tn OR quodixor:ti,ab,de,rn,tn OR resormes:ti,ab,de,rn,tn OR unomes:ti,ab,de,rn,tn OR adromux:ti,ab,de,rn,tn OR anabon:ti,ab,de,rn,tn OR bandron:ti,ab,de,rn,tn OR bantuc:ti,ab,de,rn,tn OR baxogur:ti,ab,de,rn,tn OR bonjenic:ti,ab,de,rn,tn OR bonnedra:ti,ab,de,rn,tn OR bonoste:ti,ab,de,rn,tn OR darmas:ti,ab,de,rn,tn OR disdual:ti,ab,de,rn,tn OR elasterin:ti,ab,de,rn,tn OR etanorden:ti,ab,de,rn,tn OR femorel:ti,ab,de,rn,tn OR haniban:ti,ab,de,rn,tn OR ibagenit:ti,ab,de,rn,tn OR ibames:ti,ab,de,rn,tn OR ibamyl:ti,ab,de,rn,tn OR ibandroninezuur:ti,ab,de,rn,tn OR ibandronsav:ti,ab,de,rn,tn OR ibandronsyre:ti,ab,de,rn,tn OR ibanfos:ti,ab,de,rn,tn OR ibanleg:ti,ab,de,rn,tn OR ibannate:ti,ab,de,rn,tn OR ibondro:ti,ab,de,rn,tn OR ibostofar:ti,ab,de,rn,tn OR idena:ti,ab,de,rn,tn OR ikamentin:ti,ab,de,rn,tn OR inostelid:ti,ab,de,rn,tn OR kalosso:ti,ab,de,rn,tn OR kefort:ti,ab,de,rn,tn OR licobondrat:ti,ab,de,rn,tn OR mirdezel:ti,ab,de,rn,tn OR modifical:ti,ab,de,rn,tn OR osma:ti,ab,de,rn,tn OR osteonat:ti,ab,de,rn,tn OR osteoviva:ti,ab,de,rn,tn OR phacebonate:ti,ab,de,rn,tn OR ratiban:ti,ab,de,rn,tn OR recaxin:ti,ab,de,rn,tn OR ribobandron:ti,ab,de,rn,tn OR 'r 484':ti,ab,de,rn,tn OR r484:ti,ab,de,rn,tn OR '114084 78 5':ti,ab,de,rn,tn OR '138844 81 2':ti,ab,de,rn,tn OR '138926 19 9':ti,ab,de,rn,tn	5551
#6	zoledronic AND 'acid'/exp OR (zoledronic:ti,ab,de,rn,tn AND acid:ti,ab,de,rn,tn) OR aclasta:ti,ab,de,rn,tn OR reclast:ti,ab,de,rn,tn OR 'cgp 42446':ti,ab,de,rn,tn OR cgp42446:ti,ab,de,rn,tn OR 'cgp 42446a':ti,ab,de,rn,tn OR cgp42446a:ti,ab,de,rn,tn OR orazol:ti,ab,de,rn,tn OR 'zol 446':ti,ab,de,rn,tn OR zol446:ti,ab,de,rn,tn OR blaztere:ti,ab,de,rn,tn OR bolenic:ti,ab,de,rn,tn OR boncur:ti,ab,de,rn,tn OR celdron:ti,ab,de,rn,tn OR desibon:ti,ab,de,rn,tn OR drometa:ti,ab,de,rn,tn OR eriophos:ti,ab,de,rn,tn OR fayton:ti,ab,de,rn,tn OR kaliksir:ti,ab,de,rn,tn OR ledron:ti,ab,de,rn,tn OR osporil:ti,ab,de,rn,tn OR ostezolen:ti,ab,de,rn,tn OR rionit:ti,ab,de,rn,tn OR simpla:ti,ab,de,rn,tn OR sinresor:ti,ab,de,rn,tn OR steozol:ti,ab,de,rn,tn OR synblasta:ti,ab,de,rn,tn OR syndronic:ti,ab,de,rn,tn OR varidronico:ti,ab,de,rn,tn OR zelinda:ti,ab,de,rn,tn OR zidolamin:ti,ab,de,rn,tn OR zidronic:ti,ab,de,rn,tn OR ziduvin:ti,ab,de,rn,tn OR zinvel:ti,ab,de,rn,tn OR zobone:ti,ab,de,rn,tn OR zobonic:ti,ab,de,rn,tn OR zolacitor:ti,ab,de,rn,tn OR zolako:ti,ab,de,rn,tn OR zoledro:ti,ab,de,rn,tn OR zoledreenos:ti,ab,de,rn,tn OR	17650

	zoledrin:ti,ab,de,rn,tn OR zoledronate:ti,ab,de,rn,tn OR zoledronsyre:ti,ab,de,rn,tn OR zolenat:ti,ab,de,rn,tn OR zolenic:ti,ab,de,rn,tn OR zoletalis:ti,ab,de,rn,tn OR zoletech:ti,ab,de,rn,tn OR zolira:ti,ab,de,rn,tn OR zomebon:ti,ab,de,rn,tn OR zomedron:ti,ab,de,rn,tn OR zomera:ti,ab,de,rn,tn OR zometa:ti,ab,de,rn,tn OR zomikos:ti,ab,de,rn,tn OR zuorui:ti,ab,de,rn,tn OR zyolix:ti,ab,de,rn,tn OR cenozoic:ti,ab,de,rn,tn OR desinobon:ti,ab,de,rn,tn OR indaferil:ti,ab,de,rn,tn OR midronic:ti,ab,de,rn,tn OR leuzotev:ti,ab,de,rn,tn OR tevadronic:ti,ab,de,rn,tn OR zacindate:ti,ab,de,rn,tn OR zalit:ti,ab,de,rn,tn OR zofaden:ti,ab,de,rn,tn OR zolacin:ti,ab,de,rn,tn OR zoldria:ti,ab,de,rn,tn OR zoledo:ti,ab,de,rn,tn OR zolecan:ti,ab,de,rn,tn OR zoledronsav:ti,ab,de,rn,tn OR zolenia:ti,ab,de,rn,tn OR zortila:ti,ab,de,rn,tn OR '118072 93 8':ti,ab,de,rn,tn OR '131654 46 1':ti,ab,de,rn,tn OR '165800 06 6':ti,ab,de,rn,tn OR '165800 07 7':ti,ab,de,rn,tn	
#7	'denosumab'/exp OR ((denosumab:ti,ab,de,rn,tn OR amg:ti,ab,de,rn,tn) AND 162:ti,ab,de,rn,tn) OR amg162:ti,ab,de,rn,tn OR amgiva:ti,ab,de,rn,tn OR prolia:ti,ab,de,rn,tn OR '615258 40 7':ti,ab,de,rn,tn	8833
#8	'raloxifene'/exp OR raloxifene:ti,ab,de,rn,tn OR bonmax:ti,ab,de,rn,tn OR celvista:ti,ab,de,rn,tn OR keoxifene:ti,ab,de,rn,tn OR loxar:ti,ab,de,rn,tn OR loxifen:ti,ab,de,rn,tn OR 'ly 156758':ti,ab,de,rn,tn OR ly156758:ti,ab,de,rn,tn OR ly139481:ti,ab,de,rn,tn OR 'ly 139481':ti,ab,de,rn,tn OR raxeto:ti,ab,de,rn,tn OR evista:ti,ab,de,rn,tn OR fluken:ti,ab,de,rn,tn OR gynista:ti,ab,de,rn,tn OR osteoclax:ti,ab,de,rn,tn OR osteya:ti,ab,de,rn,tn OR ostiral:ti,ab,de,rn,tn OR ralosto:ti,ab,de,rn,tn OR ronixifeno:ti,ab,de,rn,tn OR aloxif:ti,ab,de,rn,tn OR optruma:ti,ab,de,rn,tn OR oxilar:ti,ab,de,rn,tn OR raloksifen:ti,ab,de,rn,tn OR ralomeer:ti,ab,de,rn,tn OR ralopharm:ti,ab,de,rn,tn OR ralover:ti,ab,de,rn,tn OR ralox:ti,ab,de,rn,tn OR raloxa:ti,ab,de,rn,tn OR raloxibone:ti,ab,de,rn,tn OR raloxiep:ti,ab,de,rn,tn OR raloxifen:ti,ab,de,rn,tn OR raloxstar:ti,ab,de,rn,tn OR raloxten:ti,ab,de,rn,tn OR '82640 04 8':ti,ab,de,rn,tn OR '84449 90 1':ti,ab,de,rn,tn	11662
#9	'abaloparatide'/exp OR abaloparatide:ti,ab,de,rn,tn OR ba058:ti,ab,de,rn,tn OR 'ba 058':ti,ab,de,rn,tn OR 'bim 44058':ti,ab,de,rn,tn OR bim44058:ti,ab,de,rn,tn OR '247062 33 5':ti,ab,de,rn,tn	429
#10	'odanacatib'/exp OR odanacatib:ti,ab,de,rn,tn OR mk0822:ti,ab,de,rn,tn OR 'mk 0822':ti,ab,de,rn,tn OR mk822:ti,ab,de,rn,tn OR 'mk 822':ti,ab,de,rn,tn OR '603139 19 1':ti,ab,de,rn,tn	655
#11	'cathepsin k inhibitor'/exp OR 'cathepsin k inhibitor*':ti,ab	847
#12	terrosa or RGB-10 or RGB10 or movymia	37
#13	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12	53496
#14	random*:ti,ab,tn OR 'clinical trial' OR 'health care quality'/exp	5270994
#15	'animal'/de	1941963
#16	'animal experiment'/de	2553907
#17	rat:ti,ab,de OR rats:ti,ab,de OR mouse:ti,ab,de OR mice:ti,ab,de OR murine:ti,ab,de OR rodent:ti,ab,de OR rodents:ti,ab,de OR hamster:ti,ab,de OR hamsters:ti,ab,de OR pig:ti,ab,de OR pigs:ti,ab,de OR porcine:ti,ab,de OR rabbit:ti,ab,de OR rabbits:ti,ab,de OR animal:ti,ab,de OR animals:ti,ab,de OR dogs:ti,ab,de OR dog:ti,ab,de OR cats:ti,ab,de OR cow:ti,ab,de OR bovine:ti,ab,de OR sheep:ti,ab,de OR ovine:ti,ab,de OR monkey:ti,ab,de OR monkeys:ti,ab,de	7389397
#18	#15 OR #16 OR #17	7389397

#19	'human'/exp	22182792
#20	'human experiment'/de	505045
#21	#19 OR #20	22184520
#22	#18 NOT (18 AND 21)	7007175
#23	#14 NOT #22	4903057
#24	'osteoporosis'/exp	130969
#25	'metabolic bone disease'	7324
#26	'bone density'/exp	93403
#27	'fracture'/exp	325365
#28	osteoporo*:ti,ab,de OR 'osteo poro*':ti,ab,de	158284
#29	(fragil* NEAR/2 (fractur* OR break*)):ti,ab	7021
#30	(osteoporotic:ti,ab AND decalcif*:ti,ab OR patholog*:ti,ab) AND decalcif*:ti,ab OR osteopeni*:ti,ab	18374
#31	((bone:ti,ab AND mineral:ti,ab AND dens*:ti,ab OR bone:ti,ab) AND loss:ti,ab OR bone:ti,ab) AND fragil*:ti,ab	10336
#32	bmd:ti,ab	48450
#33	fractur*:ti,ab	317495
#34	(bone* NEAR/2 (density OR break* OR porosity OR porotic OR decalcif*)):ti,ab,de	108129
#35	#24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34	556755
#36	#13 AND #23 AND #35	13394
#37	**#36 AND [2018-2020]/py**	**1349**

Table 66: Summary of Cochrane searches

No.	Cochrane Search	Search hits
#1	(romosozumab or Evenity or "AMG785" or "AMG-785" or "cdp-7851" or "cdp7851" or "909395-70-6"):ti,ab,kw	98
#2	MeSH descriptor: [Teriparatide] this term only	327
#3	(Teriparatide or Forteo or Forsteo or "chs-13340" or "chs13340" or "ly-333334" or "ly333334" or parathar or "parathormone 1 34" or "parathyroid hormone 1-34" or "pth1-34" or "sun-e3001" or "sune3001" or "12583-68-5" or "52232-67-4"):ti,ab,kw	754
#4	#2 or #3	754
#5	MeSH descriptor: [Alendronate] this term only	735
#6	("alendronic acid" or alendronate or alenato or alend or alendros or alovell or arendal or bifemelan or bifosa or binosto or bonapex or defixal or dronal or endronax or eucalen or fixopan or fosalan or fosamax or fosmin or fosval or marvil or maxibone or "mk-0217" or "mk-217" or "mk0217" or "mk217" or neobon or oncalst or onclast or osdron or osdronat or oseotenk or osficar or oslene or osteofar or osteofos or osteopor or osteosan or osteovan or osticalcin or porosal or teiroc or tibolene or voroste or Fosavance or Adrovance or Vantavo or Binosto or mylan or Adronat or Alendro or Alendraccord or Alendrobell or "Alendrocor-10" or "Densate-70" or Dronalen-Plus or Ossmax or "66376-36-1"):ti,ab,kw	1840
#7	#5 or #6	1840
#8	MeSH descriptor: [Risedronate Sodium] this term only	250
#9	("risedronic acid" or actonel or atelvia or benet or "ne-58095" or "ne58095" or optinate or ribastamin or risedronate or Acris or Risedro or benet or "CO Risedrocal Combo Kit" or aktonate or bonna or cladronate or ductonar or goyart or melenor or ostenel or osteodronate or "ribastamin duo rigat" or risate or risedron or risedrogen or risendronat or risemylan or risendal or isendros or risetab or risofos or risonato or salost or tracost or acrel or actomax or actojenic or actokit or arilex or atconate or bondapen or boneact or boncur or bonmate or bontonel or bontrol or claronate or enospag or fodren or juverital or medeoros or miosen or natalox or norifax or norsed or osodens or osteoron or ostron or pexalit or tentop or resorpate or retonel or ribastamin or ribidron or ribone or richbone or ridbone or ridron or ridrone or risadican or risbon or risebon or risebone or risedon or risedreenos or risedronaat or riselib or risemed or risedrenos or risenex or risenil or riseto or risetron or resmyl or risofos or risonate or risonato or risostad or ristonat or sedron or seralis or tecnodron or tevanel or varibona or norifaz or zectoel or acridon or ridroqueen or "105462-24-6" or "122458-82-6"):ti,ab,kw	731
#10	#8 or #9	731
#11	(ibandronate or "ibandronic acid" or bonviva or bondronat or bondronate or boniva or destara or "bm-210955" or "bm210955" or bondenza or iasibon or ibandronico or alvodron or alvodronic or bandro or baxogar or bomanes or bonefrubit or bonefurbit or bonese or bonicid or bonmore or clastec or dronaval or fijical or holmevis or ibanat or ibandra or ibandrix or ibandronat or ibandronian or ibandronsav or ibanic or ibanos or ibone or ibrac or idena or ikametin or indrofar or ipexal or kefort or kemidat or licobondrat or meliba or nucodran or osagrand or osbonelle or oseum or ossica or osteocalcit or osteolong or osteosyl or ostone or posclim or quodixor or recaxin or resormes or unomes or adromux or anabon or	485

	bandron or bantuc or baxogur or bonjenic or bonnedra or bonoste or darmas or disdual or elasterin or etanorden or femorel or haniban or ibagenit or ibames or ibamyl or Ibandroninezuur or Ibandronsav or ibandronsyre or ibanfos or ibanleg or ibannate or ibondro or ibostofar or idena or ikamentin or inostelid or kalosso or kefort or licobondrat or mirdezel or modifical or osma or osteonat or osteoviva or phacebonate or ratiban or recaxin or ribobandron or "r-484" or "r484" or "114084-78-5" or "138844-81-2" or "138926-19-9"):ti,ab,kw	
#12	("zoledronic acid" or zoledronate or Aclasta or Reclast or "cgp-42446" or "cgp42446" or "cgp-42446a" or "cgp42446a" or orazol or "zol-446" or "zol446" or zomera or zometa or blaztere or bolenic or boncur or celdron or desibon or drometa or eriophos or fayton or kaliksir or ledron or osporil or ostezolen or rionit or simpla or sinresor or steozol or synblasta or syndronic or varidronico or zelinda or zidolamin or zidronic or ziduvin or zinvel or zobone or zobonic or zolacitor or zolako or zoledro or zoledreenos or zoledrin or zoledronate or zoledronsyre or zolenat or zolenic or zoletalis or zoletech or zolira or zomebon or zomedron or zomera or zometa or zomikos or zuorui or zyolix or cenozoic or desinobon or indaferil or midronic or leuzotev or tevadronic or zacindate or zalit or zofaden or zolacin or zoldria or zoledo or zolecan or zoledronsav or zolenia or zortila or "118072-93-8" or "131654-46-1" or "165800-06-6" or "165800-07-7"):ti,ab,kw	1875
#13	MeSH descriptor: [Denosumab] this term only	307
#14	(denosumab or "amg 162" or "amg162" or amgiva or prolia or "615258-40-7"):ti,ab,kw	957
#15	#13 or #14	957
#16	MeSH descriptor: [Raloxifene Hydrochloride] this term only	480
#17	(Raloxifene or "LY139481" or "LY-139481" or bonmax or celvista or evista or keoxifene or loxar or loxifen or "ly-156758" or "ly156758" or "ly139481" or "ly-139481" or raxeto or evista or fluken or gynista or osteoclax or osteya or ostiral or ralosto or raloxa or ronixifeno or aloxif or optruma or oxilar or raloksifen or ralomeer or ralopharm or ralover or ralox or raloxa or raloxibone or raloxiep or raloxifen or raloxstar or raloxten or "82640-04-8" or "84449-90-1"):ti,ab,kw	981
#18	#16 or #17	981
#19	(Abaloparatide or "BA058" or "BA-058" or "bim-44058" or "bim44058" or "247062-33-5"):ti,ab,kw	93
#20	MeSH descriptor: [Cathepsin K] this term only and with qualifier(s): [Antagonists & inhibitors - AI]	21
#21	(Odanacatib or "MK0822" or "MK-0822" or "mk822" or "mk-822" or "603139-19-1"):ti,ab,kw	105
#22	("Cathepsin K inhibitor" or "Cathepsin K inhibitors"):ti,ab	70
#23	#20 or #21 or #22	136
#24	terrosa or RGB-10 or RGB10 or movymia	6
#25	#1 or #4 or #7 or #10 or #11 or #12 or #15 or #18 or #19 or #23 or #24	6571
#26	MeSH descriptor: [Osteoporosis] explode all trees	4081
#27	MeSH descriptor: [Bone Diseases, Metabolic] this term only	515
#28	MeSH descriptor: [Bone Density] this term only	4559
#29	MeSH descriptor: [Fractures, Bone] explode all trees	5912

#30	(osteoporo* or osteo-poro*):ti,ab,kw	10982
#31	(fragil* near/2 (fractur* or break*)):ti,ab	481
#32	((osteoporotic near/2 decalcif*) or (patholog* near/2 decalcif*) or osteopeni*):ti,ab	1251
#33	((bone near/2 mineral near/2 dens*) or "bone loss" or (bone near/2 fragil*)):ti,ab,kw	10856
#34	(BMD or fractur*):ti,ab	23563
#35	#26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34	32971
#36	(bone* near/2 (density or break* or porosity or porotic or decalcif*)):ti,ab,kw	11662
#37	#26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36	34109
#38	#25 and #37	4473
#39	**#38 in Trials with publication year from 2018 to 2020, in Trials**	**484**

Appendix A2

Only trials comparing interventions within the indication of interest were included when clinical data was available.

NIH Clinicaltrials.gov

Expert search syntax option

(osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR PMO OR "pathologic decalcification" OR "pathological decalcification" OR ostepenia OR osteopenic OR "bone mineral density" OR "fragility fracture" OR "fragility fractures" OR "bone loss" OR "bone density") AND (romosozumab OR AMG785 OR "909395-70-6" OR Teriparatide OR Forteo OR "12583-68-5" OR "52232-67-4" OR Alendronate OR "alendronic acid" OR "66376-36-1" OR Risedronate OR "risedronic acid" OR actonel OR "105462-24-6" OR "122458-82-6" OR ibandronate OR "ibandronic acid" OR bonviva OR "114084-78-5" OR "138844-81-2" OR "138926-19-9" OR "zoledronic acid" OR zoledronate OR Aclasta OR Reclast OR "118072-93-8" OR "131654-46-1" OR "165800-06-6" OR "165800-07-7" OR denosumab OR "amg 162" OR "amg162" OR amgiva OR prolia OR "615258-40-7" OR Raloxifene OR bonmax OR "82640-04-8" OR "84449-90-1" OR Abaloparatide OR "247062-33-5" OR Odanacatib OR "603139-19-1")

WHO International Clinical Trials Registry Platform (ICTRP)

Advanced search option

– Recruitment status = ALL

A summary of the search strategies is provided in **Table 67**.

Table 67: WHO International Clinical Trials Registry Platform Searches

Facet 1	AND	Facet 2	Trial results
romosozumab OR AMG785 OR 909395-70-6 OR Teriparatide OR Forteo OR 12583-68-5 OR 52232-67-4 OR Alendronate OR alendronic OR 66376-36-1 OR Risedronate OR risedronic OR actonel OR 105462-24-6 OR 122458-82-6 OR ibandronate OR ibandronic OR bonviva OR	AND	osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR pathologic decalcification *(Condition)*	541

114084-78-5 OR 138844-81-2 OR 138926-19-9 OR zoledronic OR zoledronate OR Aclasta OR Reclast OR 118072-93-8 OR 131654-46-1 OR 165800-06-6 OR 165800-07-7 OR denosumab OR amg 162 OR amg162 OR amgiva OR prolia OR 615258-40-7 OR Raloxifene OR bonmax OR 82640-04-8 OR 84449-90-1 OR Abaloparatide OR 247062-33-5 OR Odanacatib OR 603139-19-1 *(Title)*			
romosozumab OR AMG785 OR 909395-70-6 OR Teriparatide OR Forteo OR 12583-68-5 OR 52232-67-4 OR Alendronate OR alendronic OR 66376-36-1 OR Risedronate OR risedronic OR actonel OR 105462-24-6 OR 122458-82-6 OR ibandronate OR ibandronic OR bonviva OR 114084-78-5 OR 138844-81-2 OR 138926-19-9 OR zoledronic OR zoledronate OR Aclasta OR Reclast OR 118072-93-8 OR 131654-46-1 OR 165800-06-6 OR 165800-07-7 OR denosumab OR amg 162 OR amg162 OR amgiva OR prolia OR 615258-40-7 OR Raloxifene OR bonmax OR 82640-04-8 OR 84449-90-1 OR Abaloparatide OR 247062-33-5 OR Odanacatib OR 603139-19-1 *(Title)*	AND	pathological decalcification OR ostepenia OR osteopenic OR bone mineral density OR fragility fracture OR fragility fractures OR bone loss OR bone density *(Condition)*	577
osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR pathologic decalcification OR pathological decalcification OR ostepenia OR osteopenic OR bone mineral density OR fragility fracture OR fragility fractures OR bone loss OR bone density *(Title)*		romosozumab OR AMG785 OR 909395-70-6 OR Teriparatide OR Forteo OR 12583-68-5 OR 52232-67-4 OR Alendronate OR alendronic OR 66376-36-1 OR Risedronate *(Intervention)*	344
osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR pathologic decalcification OR pathological decalcification OR ostepenia OR osteopenic OR bone mineral density OR fragility fracture OR fragility fractures OR bone loss OR bone density *(Title)*		risedronic OR actonel OR 105462-24-6 OR 122458-82-6 OR ibandronate OR ibandronic OR bonviva OR 114084-78-5 OR 138844-81-2 OR 138926-19-9 *(Intervention)*	76
osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR pathologic decalcification OR pathological decalcification OR ostepenia OR osteopenic OR bone mineral density OR fragility fracture OR fragility fractures OR bone loss OR bone density *(Title)*		zoledronic OR zoledronate OR Aclasta OR Reclast OR 118072-93-8 OR 131654-46-1 OR 165800-06-6 OR 165800-07-7 OR denosumab OR amg 162 OR amg162 *(Intervention)*	247

osteoporosis OR osteoporotic OR osteo-porosis OR osteo-porotic OR BMD OR pathologic decalcification OR pathological decalcification OR ostepenia OR osteopenic OR bone mineral density OR fragility fracture OR fragility fractures OR bone loss OR bone density *(Title)*		amgiva OR prolia OR 615258-40-7 OR Raloxifene OR bonmax OR 82640-04-8 OR 84449-90-1 OR Abaloparatide OR 247062-33-5 OR Odanacatib OR 603139-19-1 *(Intervention)*	154
Total (including duplicates)			1939
Total (after deduplication)			792

Appendix A3

Relevant conference publications from NOF, NOS, WCO-IOF-ESCEO were mainly captured within the used Embase search strategy as presented above. This syntax was employed to include conference abstracts and proceedings.

Another way of identifying relevant conference publications was via the Northern Light Life Sciences Conference Abstracts database. The Ovid search strategy is presented in **Table 68** below.

Northern Light Life Sciences Conference Abstracts (Ovid)

Table 68: Northern Light Life Science Conference Abstract Searches

#	Search	Hits
1	(romosozumab or Evenity or AMG785 or AMG-785 or cdp-7851 or cdp7851 or 909395-70-6).ti,ab,hw. (35)	35
2	Teriparatide/ (754)	754
3	(Teriparatide or Forteo or Forsteo or chs-13340 or chs13340 or ly-333334 or ly333334 or parathar or "parathormone 1 34" or "parathyroid hormone 1-34" or "pth[1-34]" or "sun-e3001" or "sune3001" or 12583-68-5 or 52232-67-4).ti,ab,hw. (853)	853
4	or/2-3 (853)	853
5	Alendronate/ (1191)	1,191
6	(alendronic acid or alendronate or alenato or alend or alendros or alovell or arendal or bifemelan or bifosa or binosto or bonapex or defixal or dronal or endronax or eucalen or fixopan or fosalan or fosamax or fosmin or fosval or marvil or maxibone or "mk-0217" or mk-217 or mk0217 or mk217 or neobon or oncalst or onclast or osdron or osdronat or oseotenk or osficar or oslene or osteofar or osteofos or osteopor or osteosan or osteovan or osticalcin or porosal or teiroc or tibolene or voroste or Fosavance or Adrovance or Vantavo or Binosto or mylan or Adronat or Alendro or Alendraccord or Alendrobell or Alendrocor-10 or Densate-70 or Dronalen-Plus or Ossmax or 66376-36-1).ti,ab,hw. (1250)	1,250
7	or/5-6 (1250)	1,250
8	Risedronate Sodium/ (1)	1
9	(risedronic acid or actonel or atelvia or benet or ne-58095 or ne58095 or optinate or ribastamin or risedronate or Acris or Risedro or benet or "CO Risedrocal Combo Kit" or aktonate or bonna or cladronate or ductonar or goyart or melenor or ostenel or osteodronate or "ribastamin duo rigat" or risate or risedron or risedrogen or risendronat or risemylan or risendal or isendros or risetab or risofos or risonato or	195

	salost or tracost or acrel or actomax or actojenic or actokit or arilex or atconate or bondapen or boneact or boncur or bonmate or bontonel or bontrol or claronate or enospag or fodren or juverital or medeoros or miosen or natalox or norifax or norsed or osodens or osteoron or ostron or pexalit or tentop or resorpate or retonel or ribastamin or ribidron or ribone or richbone or ridbone or ridron or ridrone or risadican or risbon or risebon or risebone or risedon or risedreenos or risedronaat or riselib or risemed or risedrenos or risenex or risenil or riseto or risetron or resmyl or risofos or risonate or risonato or risostad or ristonat or sedron or seralis or tecnodron or tevanel or varibona or norifaz or zectoel or acridon or ridroqueen or 105462-24-6 or 122458-82-6).ti,ab,hw. (195)	
10	or/8-9 (195)	195
11	(ibandronate or ibandronic acid or bonviva or bondronat or bondronate or boniva or destara or bm-210955 or bm210955 or bondenza or iasibon or ibandronico or alvodron or alvodronic or bandro or baxogar or bomanes or bonefrubit or bonefurbit or bonese or bonicid or bonmore or clastec or dronaval or fijical or holmevis or ibanat or ibandra or ibandrix or ibandronat or ibandronian or ibandronsav or ibanic or ibanos or ibone or ibrac or idena or ikametin or indrofar or ipexal or kefort or kemidat or licobondrat or meliba or nucodran or osagrand or osbonelle or oseum or ossica or osteocalcit or osteolong or osteosyl or ostone or posclim or quodixor or recaxin or resormes or unomes or adromux or anabon or bandron or bantuc or baxogur or bonjenic or bonnedra or bonoste or darmas or disdual or elasterin or etanorden or femorel or haniban or ibagenit or ibames or ibamyl or Ibandroninezuur or Ibandronsav or ibandronsyre or ibanfos or ibanleg or ibannate or ibondro or ibostofar or idena or ikamentin or inostelid or kalosso or kefort or licobondrat or mirdezel or modifical or osma or osteonat or osteoviva or phacebonate or ratiban or recaxin or ribobandron or r-484 or r484 or 114084-78-5 or 138844-81-2 or 138926-19-9).ti,ab,hw. (188)	188
12	(zoledronic acid or zoledronate or Aclasta or Reclast or cgp-42446 or cgp42446 or cgp-42446a or cgp42446a or orazol or zol-446 or zol446 or zomera or zometa or blaztere or bolenic or boncur or celdron or desibon or drometa or eriophos or fayton or kaliksir or ledron or osporil or ostezolen or rionit or simpla or sinresor or steozol or synblasta or syndronic or varidronico or zelinda or zidolamin or zidronic or ziduvin or zinvel or zobone or zobonic or zolacitor or zolako or zoledro or zoledreenos or zoledrin or zoledronate or zoledronsyre or zolenat or zolenic or zoletalis or zoletech or zolira or zomebon or zomedron or zomera or zometa or zomikos or zuorui or zyolix or cenozoic or desinobon or indaferil or midronic or leuzotev or tevadronic or zacindate or zalit or zofaden or zolacin or zoldria or zoledo or zolecan or zoledronsav or zolenia or zortila or 118072-93-8 or 131654-46-1 or 165800-06-6 or 165800-07-7).ti,ab,hw. (1135)	1,285
13	Denosumab/ (1285)	1,285
14	(denosumab or amg 162 or amg162 or amgiva or prolia or 615258-40-7).ti,ab,hw. (1286)	1,286
15	or/13-14 (1286)	1,286
16	Raloxifene Hydrochloride/ (0)	0
17	(Raloxifene or LY139481 or LY-139481 or bonmax or celvista or evista or keoxifene or loxar or loxifen or ly-156758 or ly156758 or ly139481 or ly-139481 or raxeto or evista or fluken or gynista or osteoclax or osteya or ostiral or ralosto or raloxa or ronixifeno or aloxif or optruma or oxilar or raloksifen or ralomeer or ralopharm or ralover or ralox or raloxa or raloxibone or raloxiep or raloxifen or raloxstar or raloxten or 82640-04-8 or 84449-90-1).ti,ab,hw. (274)	274
18	or/16-17 (274)	274
19	(Abaloparatide or BA058 or BA-058 or bim-44058 or bim44058 or 247062-33-5).ti,ab,hw. (48)	48
20	Cathepsin K/ai [Antagonists & Inhibitors] (0)	0

21	(Odanacatib or MK0822 or MK-0822 or mk822 or mk-822 or 603139-19-1).ti,ab,hw. (78)	78
22	"Cathepsin K inhibitor$".ti,ab,ot. (72)	72
23	or/20-22 (126)	126
24	or/1,4,7,10-12,15,18-19,23 (4402)	4,402
25	exp Osteoporosis/ (14732)	14,732
26	Bone Diseases, Metabolic/ (0)	0
27	Bone Density/ (0)	0
28	exp Fractures, Bone/ (381)	381
29	(osteoporo$ or osteo-poro$).ti,ab,ot,hw. (15709)	15,709
30	(fragil$ adj2 (fractur$ or break$)).ti,ab,ot. (670)	670
31	(osteoporotic decalcif$ or patholog$ decalcif$ or osteopeni$).ti,ab,ot. (1227)	1,227
32	(bone mineral dens$ or bone loss or bone fragil$).ti,ab,ot,hw. (6546)	6,546
33	BMD.ti,ab,ot. (5822)	5,822
34	fractur$.ti,ab,ot. (18035)	18,035
35	(bone$ adj2 (density or break$ or porosity or porotic or decalcif$)).ti,ab,ot,hw. (5416)	5,416
36	or/25-35 (34171)	34,171
37	**24 and 36 (2606)**	**2,606**

Appendix A20

UCB can confirm that all of the tables and figures in sections D.4.3, D.4.4 and D.4.5 have now been checked. In addition, all the percentages to the Events/N columns for all tables in these sections were added. Please note, Figure 16 is correct and thus corrections were not necessary. The corresponding table was corrected by removing Hadji 2012 and adding FRAME and Chao 2013.

ITT population

Table 69: New vertebral fractures – 12 months

Trial/Study	Intervention	Events/N (%)	Comparator	Events/N (%)
VERT-MN (Australia and Europe)	Risedronate	19/333 (5.7%)	Placebo	45/334 (13.5%)
VERT-MN (North America)	Risedronate	16/669 (2.4%)	Placebo	42/660 (6.4%)
FREEDOM	Denosumab	32/3702 (0.9%)	Placebo	82/3691 (2.2%)
HORIZON-PFT	Zoledronate	42/2822 (1.5%)	Placebo	106/2853 (3.7%)
Dursun et al. 2001	Alendronate	12/38 (31.6%)	Placebo	14/35 (40.0%)
Lufkin et al. 1998	Raloxifene	21/43 (48.8%)	Placebo	18/45 (40.0%)
Liu et al. 2004†	Raloxifene	0/102 (0.0%)	Placebo	5/102 (4.9%)
MORE	Raloxifene	17/2259 (0.8%)	Placebo	32/2292 (1.4%)
Morii et al. 2003†	Raloxifene	0/92 (0.0%)	Placebo	2/97(2.1%)
ARCH (ITT)	Romosozumab	82/2046 (4.0%)	Alendronate	128/2047 (6.3%)

FRAME (ITT)	Romosozumab	16/3321 (0.5%)	Placebo	59/3322 (1.8%)
VERO	Teriparatide	1/574 (0.2%)	Risedronate	35/585 (6.0%)

Footnotes: †: Zero correction was applied.
Abbreviations: ITT: Intention-to-treat.

Table 70: New vertebral fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)	Comparator	Events/N (%)
VERT-MN (Australia and Europe)	Risedronate	41/344 (11.9%)	Placebo	85/346 (24.6%)	NA	NA
ZONE	Zoledronate	9/309 (2.9%)	Placebo	27/308 (8.8%)	NA	NA
FREEDOM	Denosumab	53/3702 (1.4%)	Placebo	183/3691 (5.0%)	NA	NA
HORIZON-PFT	Zoledronate	62/2822 (2.2%)	Placebo	220/2853 (7.7%)	NA	NA
Neer et al. 2001	Teriparatide	22/444 (5.0%)	Placebo	64/448 (14.3%)	NA	NA
Hadji et al. 2012	Teriparatide	16/360 (4.4%)	Risedronate	33/350 (9.4%)	NA	NA
Bai et al. 2013	Zoledronate	6/242 (2.5%)	Placebo	9/241 (3.7%)	NA	NA
MORE	Raloxifene	105/2259 (4.6%)	Placebo	167/2292 (7.3%)	NA	NA
FIT I	Alendronate	42/955 (4.4%)	Placebo	109/940 (11.6%)	NA	NA
ARCH (ITT)	ROMO/ALN	127/2046 (6.2%)	Alendronate	243/2047 (11.9%)	NA	NA
VERO	Teriparatide	28/516 (5.4%)	Risedronate	64/533 (12.0%)	NA	NA
ACTIVE*	Abaloparatide	4/690 (0.6%)	Placebo	30/711 (4.2%)	Teriparatide	6/717 (0.8%)

Footnotes: *: Based on 18 months
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

Table 71: New vertebral fractures – 36 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
MORE	Raloxifene	148/2259 (6.6%)	Placebo	231/2292 (10.1%)
Silverman et al. 2008	Raloxifene	43/1849 (2.3%)	Placebo	77/1885 (4.1%)
VERT-MN (Australia and Europe)	Risedronate	53/344 (15.4%)	Placebo	89/346 (25.7%)
VERT-MN (North America)	Risedronate	61/696 (8.8%)	Placebo	93/678 (13.7%)
FREEDOM	Denosumab	86/370 (2.3%)	Placebo	264/3691 (7.2%)
HORIZON-PFT	Zoledronate	92/2822 (3.3%)	Placebo	310/2853 (10.9%0

FIT I	Alendronate	78/981 (8.0%)	Placebo	145/965 (15.0%)
Liberman et al. 1995	Alendronate	5/196 (2.6%)	Placebo	22/355 (6.2%)
ARCH (ITT)*	ROMO/ALN	88/1833 (4.8%)	Alendronate	178/1838 (9.7%)

Footnotes: *36 month CSR data used for ITT population.
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

Table 72: Non-vertebral fractures – 12 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FOSIT	Alendronate	19/950 (2%)	Placebo	37/958 (3.9%)
VERT-MN (Europe)	Risedronate	19/406 (4.7%)	Placebo	23/406 (5.7%)
Neer et al. 2001	Teriparatide	23/541 (4.3%)	Placebo	29/544 (5.3%)
FREEDOM	Denosumab	101/3902 (2.6%)	Placebo	121/3906 (3.1%)
CHAO	Zoledronate	5/327 (1.5%)	Placebo	14/333 (4.2%)
HORIZON	Zoledronate	126/3875 (3.3%)	Placebo	149/3861 (3.9%)
RUTH	Raloxifene	70/5044 (1.4%)	Placebo	60/5057 (1.2%)
VERO	Teriparatide	15/680 (2.2%)	Risedronate	21/680 (3.1%)
FRAME (ITT)	Romosozumab	56/3589 (1.6%)	Placebo	75/3591 (2.1%)
ARCH (ITT)	Romosozumab	70/2046 (3.4%)	Alendronate	95/2047 (4.6%)

Abbreviations: ITT: Intention-to-treat.

Table 73: Non-vertebral fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)	Comparator	Events/N (%)
ACTIVE ‡	Abaloparatide	18/824 (2.2%)	Placebo	33/821 (4%)	Teriparatide	24/818 (2.9%)
VERO	Teriparatide	25/680 (3.7%)	Risedronate	38/680 (5.6%)	NA	NA
VERT-MN (Europe)	Risedronate	29/406 (7.1%)	Placebo	46/406 (11.3%)	NA	NA
Hadji et al. 2012	Teriparatide	28/360 (7.8%)	Risedronate	29/350 (8.3%)	NA	NA
Neer et al. 2001	Teriparatide	34/541 (6.3%)	Placebo	53/544 (9.7%)	NA	NA
FREEDOM	Denosumab	179/3902 (4.6%)	Placebo	227/3906 (5.8%)	NA	NA
HORIZON	Zoledronate	221/3875 (5.7%)	Placebo	281/3861 (7.3%)	NA	NA
RUTH	Raloxifene	142/5044 (2.8%)	Placebo	150/5057 (3%)	NA	NA

FOSIT *	Alendronate	19/950 (2%)	Placebo	37/958 (3.9%)	NA	NA
ARCH (ITT)	ROMO/ALN	129/2046 (6.3%)	Alendronate	159/2047 (7.8%)	NA	NA

Footnotes: ‡: Based on 18 months; *: Based on 12 months to ensure the link between ROMO/ALN and placebo via alendronate.
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

Table 74: Non-vertebral fractures – 36 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
VERT-MN (Europe)	Risedronate	36/406 (8.9%)	Placebo	51/406 (12.6%)
VERT-MN (North America)	Risedronate	33/812 (4.1%)	Placebo	52/815 (6.4%)
FREEDOM	Denosumab	238/3902 (6.1%)	Placebo	293/3906 (7.5%)
HORIZON	Zoledronate	292/3875 (7.5%)	Placebo	388/3861 (10%)
FIT I+II	Alendronate	283/3236 (8.7%)	Placebo	442/3223 (13.7%)
Chao et al. 2013	Zoledronate	28/327 (8.6%)	Placebo	48/333 (14.4%)
RUTH	Raloxifene	213/5044 (4.2%)	Placebo	214/5057 (4.2%)
Silverman et al. 2008	Raloxifene	109/1849 (5.9%)	Placebo	119/1885 (6.3%)
ARCH (ITT) *	ROMO/ALN	178/2046 (8.7%)	Alendronate	217/2047 (10.6%)

Footnotes: *Based on primary analysis data 30 months.
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

Table 75: Hip fractures – 12 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FREEDOM	Denosumab	12/3902 (0.3%)	Placebo	23/3906 (0.6%)
HORIZON-PFT	Zoledronate	24/3875 (0.6%)	Placebo	31/3861 (0.8%)
Chao et al. 2013	Zoledronate	3/327 (0.9%)	Placebo	4/333 (1.2%)
RUTH	Raloxifene	13/5044 (0.3%)	Placebo	8/5057 (0.2%)
ARCH (ITT)	Romosozumab	14/2046 (0.7%)	Alendronate	22/2047 (1.1%)
FRAME (ITT)	Romosozumab	7/3589 (0.2%)	Placebo	13/3591 (0.4%)
EVA‡	Alendronate	1/716 (0.1%)	Placebo	2/707 (0.3%)

Footnotes: ‡: Based on 10.3 months.
Abbreviations: ITT: Intention-to-treat.

Table 76: Hip fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)	Comparator	Events/N (%)
Hadji et al. 2012	Teriparatide	5/360 (1.4%)	Risedronate	2/350 (0.6%)	NA	NA
VERO	Teriparatide	2/680 (0.3%)	Risedronate	5/680 (0.7%)	NA	NA
Neer et al. 2001	Teriparatide	2/541 (0.4%)	Placebo	4/544 (0.7%)	NA	NA
ACTIVE†	Abaloparatide	0/824 (0%)	Teriparatide	0/818 (0%)	Placebo	2/821 (0.2%)
FREEDOM	Denosumab	16/3902 (0.4%)	Placebo	35/3906 (0.9%)	NA	NA
Bai et al. 2013	Zoledronate	12/242 (5%)	Placebo	21/241 (8.7%)	NA	NA
HORIZON-PFT	Zoledronate	40/3875 (1%)	Placebo	56/3861 (1.5%)	NA	NA
RUTH	Raloxifene	21/5044 (0.4%)	Placebo	23/5057 (0.5%)	NA	NA
FIT I‡	Alendronate	11/1022 (1.1%)	Placebo	22/1005 (2.2%)	NA	NA
ARCH (ITT)	ROMO/ALN	31/2046 (1.5%)	Alendronate	43/2047 (2.1%)	NA	NA

Footnotes: †: Based on 18 months, zero correction was applied; ‡: Based on 36 months to ensure the link between ROMO/ALN and placebo via alendronate.
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

Table 77: Hip fractures – 36 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FREEDOM	Denosumab	26/3902 (0.7%)	Placebo	92/3906 (2.4%)
HORIZON	Zoledronate	52/3875 (1.3%)	Placebo	88/3861 (2.3%)
FIT I	Alendronate	11/1022 (1.1%)	Placebo	22/1005 (2.2%)
ARCH (ITT)*	ROMO/ALN	41/2046 (2%)	Alendronate	66/2047 (3.2%)
VERT-MN (Europe)	Risedronate	9/406 (2.2%)	Placebo	11/406 (2.7%)
VERT-MN (North America)	Risedronate	12/812 (1.5%)	Placebo	15/815 (1.8%)
CHAO	Zoledronate	8/327 (2.4%)	Placebo	13/333 (3.9%)
RUTH	Raloxifene	35/5044 (0.7%)	Placebo	37/5057 (0.7%)

Footnotes: *Based on primary analysis data 30 months
Abbreviations: ITT: Intention-to-treat; ROMO/ALN: Romosozumab followed by alendronate.

EU LABEL population

New vertebral fractures – 12 months

Table 78: New vertebral fractures – 12 months

Trial/Study	Intervention	Events/N (%)	Comparator	Events/N (%)
VERT-MN (Australia and Europe)	Risedronate	19/333 (5.7%)	Placebo	45/334 (13.5%)
VERT-MN (North America)	Risedronate	16/669 (2.4%)	Placebo	42/660 (6.4%)

FREEDOM	Denosumab	32/3702 (0.9%)	Placebo	82/3691 (2.2%)
HORIZON-PFT	Zoledronate	42/2822 (1.5%)	Placebo	106/2853 (3.7%)
Dursun et al. 2001	Alendronate	12/38 (31.6%)	Placebo	14/35 (40%)
Lufkin et al. 1998	Raloxifene	21/43 (48.8%)	Placebo	18/45 (40%)
Liu et al. 2004†	Raloxifene	0/102 (0%)	Placebo	5/102 (4.9%)
MORE	Raloxifene	17/2259 (0.8%)	Placebo	32/2292 (1.4%)
Morii et al. 2003†	Raloxifene	0/92 (0%)	Placebo	2/97 (2.1%)
ARCH (EU)‡	Romosozumab	50/1592 (3.1%)	Alendronate	78/1598 (4.9%)
FRAME (EU)‡	Romosozumab	11/1242 (0.9%)	Placebo	28/1262 (2.2%)
VERO	Teriparatide	1/574 (0.2%)	Risedronate	35/585 (6%)

Footnotes: †: Zero correction was applied
Abbreviations: EU: European-label.

Table 79: New vertebral fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N	Comparator	Events/N (%)
VERT-MN (Australia and Europe)	Risedronate	41/344 (11.9%)	Placebo	85/346 (24.6%)	NA	NA
ZONE	Zoledronate	9/309 (2.9%)	Placebo	27/308 (8.8%)	NA	NA
FREEDOM	Denosumab	53/3702 (1.4%)	Placebo	183/3691 (5%)	NA	NA
HORIZON-PFT	Zoledronate	62/2822 (2.2%)	Placebo	220/2853 (7.7%)	NA	NA
Neer et al. 2001	Teriparatide	22/444 (5%)	Placebo	64/448 (14.3%)	NA	NA
Hadji et al. 2012	Teriparatide	16/360 (4.4%)	Risedronate	33/350 (9.4%)	NA	NA
Bai et al. 2013	Zoledronate	6/242 (2.5%)	Placebo	9/241 (3.7%)	NA	NA
MORE	Raloxifene	105/2259 (4.6%)	Placebo	167/2292 (7.3%)	NA	NA
FIT I	Alendronate	42/955 (4.4%)	Placebo	109/940 (11.6%)	NA	NA
ARCH (EU)	ROMO/ALN	69/1715 (4%)	Alendronate	133/1724 (7.7%)	NA	NA
VERO	Teriparatide	28/516 (5.4%)	Risedronate	64/533 (12%)	NA	NA
ACTIVE*	Abaloparatide	4/690 (0.6%)	Placebo	30/711 (4.2%)	Teriparatide	6/717 (0.8%)

Footnotes: * Based on 18 months

Table 80: Non-vertebral fractures – 12 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FOSIT	Alendronate	19/950 (2%)	Placebo	37/958 (3.9%)
VERT-MN (Europe)	Risedronate	19/406 (4.7%)	Placebo	23/406 (5.7%)
Neer et al. 2001	Teriparatide	23/541 (4.3%)	Placebo	29/544 (5.3%)
FREEDOM	Denosumab	101/3902 (2.6%)	Placebo	121/3906 (3.1%)
CHAO	Zoledronate	5/327 (1.5%)	Placebo	14/333 (4.2%)
HORIZON	Zoledronate	126/3877 (3.2%)	Placebo	149/3861 (3.9%)
RUTH	Raloxifene	70/5044 (1.4%)	Placebo	60/5057 (1.2%)
VERO	Teriparatide	15/680 (2.2%)	Risedronate	21/680 (3.1%)
FRAME (EU-LABEL)	Romosozumab	29/1353 (2.1%)	Placebo	37/1383 (2.7%)
ARCH (EU-LABEL)	Romosozumab	67/1923 (3.5%)	Alendronate	91/1920 (4.7%)

Table 81: Non-vertebral fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)	Comparator	Events/N (%)
ACTIVE ‡	Abaloparatide	18/824 (2.2%)	Placebo	33/821 (4%)	Teriparatide	24/818 (2.9%)
VERO	Teriparatide	25/680 (3.7%)	Risedronate	38/680 (5.6%)	NA	NA
VERT-MN (Europe)	Risedronate	29/406 (7.1%)	Placebo	46/406 (11.3%)	NA	NA
Hadji et al. 2012	Teriparatide	28/360 (7.8%)	Risedronate	29/350 (8.3%)	NA	NA
Neer et al. 2001	Teriparatide	34/541 (6.3%)	Placebo	53/544 (9.7%)	NA	NA
FREEDOM	Denosumab	179/3902 (4.6%)	Placebo	227/3906 (5.8%)	NA	NA
HORIZON	Zoledronate	221/3875 (5.7%)	Placebo	281/3861 (7.3%)	NA	NA
RUTH	Raloxifene	142/5044 (2.8%)	Placebo	150/5057 (3%)	NA	NA
FOSIT*	Alendronate	19/950 (2%)	Placebo	37/958 (3.9%)	NA	NA
ARCH (EU)**	ROMO/ALN	124/1923 (6.4%)	Alendronate	151/1920 (7.9%)	NA	NA

Footnotes: ‡ Based on 18 months; * Based on 12 months to ensure the link between ROMO/ALN and placebo via alendronate; **CSR data (post-hoc analysis).
Abbreviations: EU: European label; NA: Not applicable; ROMO/ALN: Romosozumab followed by alendronate.

Table 82: Non-vertebral fractures – 36 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
VERT-MN (Europe)	Risedronate	36/406 (8.9%)	Placebo	51/406 (12.6%)
VERT-MN (North America)	Risedronate	33/812 (4.1%)	Placebo	52/815 (6.4%)
FREEDOM	Denosumab	238/3902 (6.1%)	Placebo	293/3906 (7.5%)
HORIZON	Zoledronate	292/3875 (7.5%)	Placebo	388/3861 (10%)
FIT I+II	Alendronate	283/3236 (8.7%)	Placebo	442/3223 (13.7%)
Chao et al. 2013	Zoledronate	28/327 (8.6%)	Placebo	48/333 (14.4%)
RUTH	Raloxifene	213/5044 (4.2%)	Placebo	214/5057 (4.2%)
Silverman et al. 2008	Raloxifene	109/1849 (5.9%)	Placebo	119/1885 (6.3%)
ARCH (EU)*	ROMO/ALN	171/1923 (8.9%)	Alendronate	207/1920 (10.8%)

Footnotes: * 36 month CSR data (post-hoc analysis).
Abbreviations: ROMO/ALN: Romosozumab followed by alendronate.

Table 83: Hip fractures – 12 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FREEDOM	Denosumab	12/3902 (0.3%)	Placebo	23/3906 (0.6%)
HORIZON-PFT	Zoledronate	24/3875 (0.6%)	Placebo	31/3861 (0.8%)
Chao et al. 2013	Zoledronate	3/327 (0.9%)	Placebo	4/333 (1.2%)
RUTH	Raloxifene	13/5044 (0.3%)	Placebo	8/5057 (0.2%)
ARCH (EU)	Romosozumab	14/1923 (0.7%)	Alendronate	22/1920 (1.1%)
FRAME (EU)	Romosozumab	3/1353 (0.2%)	Placebo	9/1383 (0.7%)
EVA‡	Alendronate	1/716 (0.1%)	Placebo	2/707 (0.3%)

Footnotes: ‡ Based on 10.3 months.

Table 84: Hip fractures – 24 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)	Comparator	Events/N (%)
Hadji et al. 2012	Teriparatide	5/360 (1.4%)	Risedronate	2/350 (0.6%)	NA	NA

VERO	Teriparatide	2/680 (0.3%)	Risedronate	5/680 (0.7%)	NA	NA
Neer et al. 2001	Teriparatide	2/541 (0.4%)	Placebo	4/544 (0.7%)	NA	NA
ACTIVE†	Abaloparatide	0/824 (0%)	Teriparatide	0/818 (0%)	Placebo	2/821 (0.2%)
FREEDOM	Denosumab	16/3902 (0.4%)	Placebo	35/3906 (0.9%)	NA	NA
Bai et al. 2013	Zoledronate	12/242 (5%)	Placebo	21/241 (8.7%)	NA	NA
HORIZON-PFT	Zoledronate	40/3875 (1%)	Placebo	56/3861 (1.5%)	NA	NA
RUTH	Raloxifene	21/5044 (0.4%)	Placebo	23/5057 (0.5%)	NA	NA
FIT I‡	Alendronate	11/1022 (1.1%)	Placebo	22/1005 (2.2%)	NA	NA
ARCH (EU)	ROMO/ALN	30/1923 (1.6%)	Alendronate	42/1920 (2.2%)	NA	NA

Footnotes: † Based on 18 months, zero correction was applied; ‡ Based on 36 months.
Abbreviations: NA: Not applicable; ROMO/ALN: Romosozumab followed by alendronate.

Table 85: Hip fractures – 36 months

Trial	Intervention	Events/N (%)	Comparator	Events/N (%)
FREEDOM	Denosumab	26/3902 (0.7%)	Placebo	92/3906 (2.4%)
HORIZON	Zoledronate	52/3875 (1.3%)	Placebo	88/3861 (2.3%)
FIT I	Alendronate	11/1022 (1.1%)	Placebo	22/1005 (2.2%)
ARCH (EU)*	ROMO/ALN	40/1923 (2.1%)	Alendronate	64/1920 (3.3%)
TVERT-MN (Europe)	Risedronate	9/406 (2.2%)	Placebo	11/406 (2.7%)
VERT-MN (North America)	Risedronate	12/812 (1.5%)	Placebo	15/815 (1.8%)
CHAO	Zoledronate	8/327 (2.4%)	Placebo	13/333 (3.9%)
RUTH	Raloxifene	35/5044 (0.7%)	Placebo	37/5057 (0.7%)

Footnotes: *36-month CSR data (post-hoc analysis).
Abbreviations: ROMO/ALN: Romosozumab followed by alendronate.

Appendix A22

Fractures Fixed effects

```
# Binomial likelihood, logit link
# Fixed effects model
model{                    # *** PROGRAM STARTS
```

```
for(i in 1:NS){                # LOOP THROUGH STUDIES
   mu[i] ~ dnorm(0,.0001)      # vague priors for all trial baselines
   for (k in 1:na[i])  {       # LOOP THROUGH ARMS
      r[i,k] ~ dbin(p[i,k],n[i,k])    # binomial likelihood
# model for linear predictor
      logit(p[i,k]) <- mu[i] + d[t[i,k]] - d[t[i,1]]
# expected value of the numerators
      rhat[i,k] <- p[i,k] * n[i,k]
#Deviance contribution
      dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
          +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
    }
# summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
    }
totresdev <- sum(resdev[])      # Total Residual Deviance
d[1]<-0    # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:NT){  d[k] ~ dnorm(0,.0001) }

for (i in 1:NS) {mu1[i]<-mu[i] * equals(t[i,1],1)
nxmu1[i]<-n[i,1]*mu1[i]
n1[i]<-n[i,1]*equals(t[i,1],1) }

# ranking
for (k in 1:NT) { rk[k] <- rank(d[],k)
                  best[k]<-equals(rk[k],1)
for (h in 1:NT){ prob[h,k] <- equals(rk[k],h) }}

for (k in 1:NT)   { logit(T[k])<- sum(nxmu1[])/sum(n1[])+d[k] }
# pairwise ORs
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
             {  lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
             }
```

```
        }

# pairwise RRs
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
              { rr[c,k] <-(T[k]/T[c])
              }
        }
}                                       # *** PROGRAM ENDS
```

Fractures Random effects

```
#Random effects model for multi-arm trials (any number of arms)

model{
for(i in 1:NS){
        w[i,1] <-0
              delta[i,t[i,1]]<-0
              mu[i] ~ dnorm(0,.0001)                                    # vague priors for trial
baselines
              for (k in 1:na[i])  {
                    r[i,k] ~ dbin(p[i,t[i,k]],n[i,k])                                  # binomial likelihood
                                logit(p[i,t[i,k]])<-mu[i] + delta[i,t[i,k]] }
# model
                     for (k in 2:na[i]) {
              delta[i,t[i,k]] ~ dnorm(md[i,t[i,k]],taud[i,t[i,k]])          # trial-specific LOR distributions
              md[i,t[i,k]] <-  d[t[i,k]] - d[t[i,1]]  + sw[i,k]             # mean of LOR distributions
               taud[i,t[i,k]] <- tau *2*(k-1)/k                             #precision of LOR distributions
               w[i,k] <- (delta[i,t[i,k]]  - d[t[i,k]] + d[t[i,1]])         #adjustment, multi-arm RCTs
               sw[i,k] <-sum(w[i,1:k-1])/(k-1) }              # cumulative adjustment for multi-arm trials
 }

d[1]<-0
for (k in 2:NT){d[k] ~ dnorm(0,.0001) }                 #  vague priors for basic parameters

sd~dunif(0,0.6)                                 #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)
```

```
for (i in 1:NS) {mu1[i]<-mu[i] * equals(t[i,1],1)
nxmu1[i]<-n[i,1]*mu1[i]
n1[i]<-n[i,1]*equals(t[i,1],1) }

for (k in 1:NT)   { logit(T[k])<- sum(nxmu1[])/sum(n1[])+d[k] }

# ranking
for (k in 1:NT) { rk[k] <- rank(d[],k)
                    best[k]<-equals(rk[k],1)
for (h in 1:NT){ prob[h,k] <- equals(rk[k],h) }}

# pairwise ORs
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
              {  lor[c,k] <- d[k] - d[c]
                 log(or[c,k]) <- lor[c,k]
              }
         }

# pairwise RRs
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
              { rr[c,k] <-(T[k]/T[c])
              }
         }
}
```

BMD (random effects)

```
# Normal likelihood, identity link
# Arm and Trial-level data (treatment differences)
# Random effects model for multi-arm trials
model{                          # *** PROGRAM STARTS
for(i in 1:ns.a){                # LOOP THROUGH STUDIES WITH ARM DATA
```

```
    w.a[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
    delta[i,1] <- 0            # treatment effect is zero for control arm
    mu[i] ~ dnorm(0,.0001)          # vague priors for all trial baselines
    for (k in 1:na.a[i]) {            #  LOOP THROUGH ARMS
        var.a[i,k] <- pow(se.a[i,k],2)   # calculate variances
        prec.a[i,k] <- 1/var.a[i,k]      # set precisions
        y.a[i,k] ~ dnorm(theta[i,k],prec.a[i,k]) # normal likelihood
        theta[i,k] <- mu[i] + delta[i,k]  # model for linear predictor
#Deviance contribution
        dev[i,k] <- (y.a[i,k]-theta[i,k])*(y.a[i,k]-theta[i,k])*prec.a[i,k]
      }
#  summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na.a[i]])
    for (k in 2:na.a[i]) {            # LOOP THROUGH ARMS
# trial-specific LOR distributions
        delta[i,k] ~ dnorm(md[i,k],taud.a[i,k])
# mean of LOR distributions, with multi-arm trial correction
        md[i,k] <-  d[t.a[i,k]] - d[t.a[i,1]] + sw.a[i,k]
# precision of LOR distributions (with multi-arm trial correction)
        taud.a[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
        w.a[i,k] <- (delta[i,k] - d[t.a[i,k]] + d[t.a[i,1]])
# cumulative adjustment for multi-arm trials
        sw.a[i,k] <- sum(w.a[i,1:k-1])/(k-1)
      }
  }
for(i in 1:ns.t){        # LOOP THROUGH STUDIES WITH TRIAL DATA
    w[i,1] <- 0  # adjustment for multi-arm trials is zero for control arm
    delta[i+ns.a,1] <- 0          # treatment effect is zero for control arm
    for (k in 2:na[i]) {          #  LOOP THROUGH ARMS
        var[i,k] <- pow(se[i,k],2)   # calculate variances
        prec[i,k] <- 1/var[i,k]      # set precisions
        y[i,k] ~ dnorm(delta[i+ns.a,k],prec[i,k]) # normal likelihood
#Deviance contribution
        dev[i+ns.a,k] <- (y[i,k]-delta[i+ns.a,k])*
                    (y[i,k]-delta[i+ns.a,k])* prec[i,k]
```

```
      }
#  summed residual deviance contribution for this trial
    resdev[i+ns.a] <- sum(dev[i+ns.a,2:na[i]])
    for (k in 2:na[i]) {             # LOOP THROUGH ARMS
# trial-specific LOR distributions
        delta[i+ns.a,k] ~ dnorm(md[i+ns.a,k],taud[i,k])
# mean of LOR distributions, with multi-arm trial correction
        md[i+ns.a,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LOR distributions (with multi-arm trial correction)
        taud[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
        w[i,k] <- (delta[i+ns.a,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
        sw[i,k] <- sum(w[i,1:k-1])/(k-1)
      }
  }
totresdev <- sum(resdev[])            #Total Residual Deviance
d[1]<-0       # treatment effect is zero for control arm
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
#all mean differences of all possible comparisons
for (c in 1:(nt-1)) {
for (k in (c + 1):nt) {
pw.diff[c,k] <-(d[k] - d[c]) # pairwise differences
                                }
                        }

# ranking on relative scale
for (k in 1:nt) {
 rk[k] <- nt+1-rank(d[],k) # assumes events are “good”
# rk[k] <- rank(d[],k) # assumes events are “bad”
 best[k] <- equals(rk[k],1) #calculate probability that treat k is best
 for (h in 1:nt){ prob[h,k] <- equals(rk[k],h) } # calculates probability that treat k is h-th best
 }
```

} # *** PROGRAM ENDS

ITT input data

New vertebral fractures 12 months

list(NT=8,NS=12)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
45	334	19	333	NA	NA	1	2	NA	2	#VERT-MN (Australia and Europe)
42	660	16	669	NA	NA	1	2	NA	2	#VERT-MN (North America)
82	3691	32	3702	NA	NA	1	4	NA	2	#FREEDOM
106	2853	42	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
14	35	12	38	NA	NA	1	6	NA	2	#Dursun 2001
18	45	21	43	NA	NA	1	7	NA	2	#Lufkin 1998
5.5	103	0.5	103	NA	NA	1	7	NA	2	#Liu 2004
32	2292	17	2259	NA	NA	1	7	NA	2	#MORE
2.5	98	0.5	93	NA	NA	1	7	NA	2	#Morii 2003
128	2047	82	2046	NA	NA	6	8	NA	2	#ARCH
59	3322	16	3321	NA	NA	1	8	NA	2	#FRAME
35	585	18	574	NA	NA	2	3	NA	2	#VERO

New vertebral fractures 24 months

list(NT=9,NS=12)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
85	346	41	344	NA	NA	1	2	NA	2	#VERT-MN (Australia and Europe)
27	308	9	309	NA	NA	1	5	NA	2	#ZONE
183	3691	53	3702	NA	NA	1	4	NA	2	#FREEDOM
220	2853	62	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
64	448	22	444	NA	NA	1	3	NA	2	#Neer et al. 2001
33	350	16	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
9	241	6	242	NA	NA	1	5	NA	2	#Bai et al. 2013
167	2292	105	2259	NA	NA	1	7	NA	2	#MORE

109	940	42	955	NA	NA	1	6	NA	2	#FIT I
64	533	28	516	NA	NA	2	3	NA	2	#VERO
243	2047	127	2046	NA	NA	6	8	NA	2	#ARCH
30	711	6	717	4	690	1	3	9	3	#ACTIVE

New vertebral fractures 36 months

list(NT=7,NS=9)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
231	2292	148	2259	NA	NA	1	6	NA	2	#MORE
77	1885	43	1849	NA	NA	1	6	NA	2	#Silverman et al. 2008
89	346	53	344	NA	NA	1	2	NA	2	#VERT-MN (AUS and EU)
93	678	61	696	NA	NA	1	2	NA	2	#VERT-MN (NA)
264	3691	86	3702	NA	NA	1	4	NA	2	#FREEDOM
310	2853	92	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
145	965	78	981	NA	NA	1	3	NA	2	#FIT I
22	355	5	196	NA	NA	1	3	NA	2	#Liberman et al. 1995
178	1838	88	1833	NA	NA	3	7	NA	2	#ARCH

Non-vertebral fractures 12 months

list(NT=8,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
37	958	19	950	NA	NA	1	6	NA	2	#FOSIT
23	406	19	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
29	544	23	541	NA	NA	1	3	NA	2	#Neer et al. 2001
121	3906	101	3902	NA	NA	1	4	NA	2	#FREEDOM
14	333	5	327	NA	NA	1	5	NA	2	#CHAO
149	3861	126	3877	NA	NA	1	5	NA	2	#HORIZON
60	5057	70	5044	NA	NA	1	7	NA	2	#RUTH
21	680	15	680	NA	NA	2	3	NA	2	#VERO
75	3591	56	3589	NA	NA	1	8	NA	2	#FRAME (ITT)
95	2047	70	2046	NA	NA	6	8	NA	2	#ARCH (ITT)

Non-vertebral fractures 24 months

list(NT=9,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
33	821	24	818	18	824	1	3	9	3	#ACTIVE
38	680	25	680	NA	NA	2	3	NA	2	#VERO
46	406	29	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
29	350	28	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
53	544	34	541	NA	NA	1	3	NA	2	#Neer et al. 2001
227	3906	179	3902	NA	NA	1	4	NA	2	#FREEDOM
281	3861	221	3875	NA	NA	1	5	NA	2	#HORIZON
150	5057	142	5044	NA	NA	1	7	NA	2	#RUTH
37	958	19	950	NA	NA	1	6	NA	2	#FOSIT
159	2047	129	2046	NA	NA	6	8	NA	2	#ARCH (ITT)

Non-vertebral fractures 36 months

list(NT=7,NS=9)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
51	406	36	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
52	815	33	812	NA	NA	1	2	NA	2	#VERT-MN (North America)
293	3906	238	3902	NA	NA	1	4	NA	2	#FREEDOM
388	3861	292	3875	NA	NA	1	5	NA	2	#HORIZON
442	3223	283	3236	NA	NA	1	3	NA	2	#FIT I+II
48	333	28	327	NA	NA	1	5	NA	2	#Chao et al. 2013
214	5057	213	5044	NA	NA	1	6	NA	2	#RUTH
119	1885	109	1849	NA	NA	1	6	NA	2	#Silverman et al. 2008
217	2047	178	2046	NA	NA	3	7	NA	2	#ARCH (ITT)

Hip fractures 12 months

list(NT=6,NS=7)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
23	3906	12	3902	NA	NA	1	2	NA	2	#FREEDOM
31	3861	24	3875	NA	NA	1	3	NA	2	#HORIZON-PFT
4	333	3	327	NA	NA	1	3	NA	2	#Chao et al. 2013
8	5057	13	5044	NA	NA	1	4	NA	2	#RUTH
22	2046	14	2046	NA	NA	5	6	NA	2	#ARCH (ITT)
13	3591	7	3589	NA	NA	1	6	NA	2	#FRAME (ITT)
2	707	1	716	NA	NA	1	5	NA	2	#EVA

Hip fractures 24 months

list(NT=9,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
2	350	5	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
5	680	2	680	NA	NA	2	3	NA	2	#VERO
4	544	2	541	NA	NA	1	3	NA	2	#Neer et al. 2001
0.5	819	2.5	822	0.5	825	3	1	9	3	#ACTIVE
35	3906	16	3902	NA	NA	1	4	NA	2	#FREEDOM
21	241	12	242	NA	NA	1	5	NA	2	#Bai et al. 2013
56	3861	40	3875	NA	NA	1	5	NA	2	#HORIZON-PFT
23	5057	21	5044	NA	NA	1	7	NA	2	#RUTH
22	1005	11	1022	NA	NA	1	6	NA	2	#FIT I
43	2047	31	2046	NA	NA	6	8	NA	2	#ARCH (ITT)

Hip fractures 36 months

list(NT=7,NS=8)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
92	3906	26	3902	NA	NA	1	4	NA	2	#FREEDOM
88	3861	52	3875	NA	NA	1	5	NA	2	#HORIZON
22	1005	11	1022	NA	NA	1	3	NA	2	#FIT I
66	2047	41	2046	NA	NA	3	7	NA	2	#ARCH (ITT)
11	406	9	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
15	815	12	812	NA	NA	1	2	NA	2	#VERT-MN (North America)
13	333	8	327	NA	NA	1	5	NA	2	#CHAO

37	5057	35	5044	NA	NA	1	6	NA	2	#RUTH

EU-label input data

New vertebral fractures 12 months

list(NT=8,NS=12)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
45	334	19	333	NA	NA	1	2	NA	2	#VERT-MN (Australia and Europe)
42	660	16	669	NA	NA	1	2	NA	2	#VERT-MN (North America)
82	3691	32	3702	NA	NA	1	4	NA	2	#FREEDOM
106	2853	42	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
14	35	12	38	NA	NA	1	6	NA	2	#Dursun 2001
18	45	21	43	NA	NA	1	7	NA	2	#Lufkin 1998
5.5	103	0.5	103	NA	NA	1	7	NA	2	#Liu 2004
32	2292	17	2259	NA	NA	1	7	NA	2	#MORE
2.5	98	0.5	93	NA	NA	1	7	NA	2	#Morii 2003
78	1598	50	1592	NA	NA	6	8	NA	2	#ARCH
28	1262	11	1242	NA	NA	1	8	NA	2	#FRAME
35	585	18	574	NA	NA	2	3	NA	2	#VERO

New vertebral fractures 24 months

list(NT=9,NS=12)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
85	346	41	344	NA	NA	1	2	NA	2	#VERT-MN (Australia and Europe)
27	308	9	309	NA	NA	1	5	NA	2	#ZONE
183	3691	53	3702	NA	NA	1	4	NA	2	#FREEDOM
220	2853	62	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
64	448	22	444	NA	NA	1	3	NA	2	#Neer et al. 2001
33	350	16	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
9	241	6	242	NA	NA	1	5	NA	2	#Bai et al. 2013
167	2292	105	2259	NA	NA	1	7	NA	2	#MORE
109	940	42	955	NA	NA	1	6	NA	2	#FIT I

64	533	28	516	NA	NA	2	3	NA	2	#VERO
133	1724	69	1715	NA	NA	6	8	NA	2	#ARCH
30	711	6	717	4	690	1	3	9	3	#ACTIVE

New vertebral fractures 36 months

list(NT=7,NS=9)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
231	2292	148	2259	NA	NA	1	6	NA	2	#MORE
77	1885	43	1849	NA	NA	1	6	NA	2	#Silverman et al. 2008
89	346	53	344	NA	NA	1	2	NA	2	#VERT-MN (AUS and EU)
93	678	61	696	NA	NA	1	2	NA	2	#VERT-MN (NA)
264	3691	86	3702	NA	NA	1	4	NA	2	#FREEDOM
310	2853	92	2822	NA	NA	1	5	NA	2	#HORIZON-PFT
145	965	78	981	NA	NA	1	3	NA	2	#FIT I
22	355	5	196	NA	NA	1	3	NA	2	#Liberman et al. 1995
178	1838	88	1833	NA	NA	3	7	NA	2	#ARCH

Non-vertebral fractures 12 months

list(NT=8,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
37	958	19	950	NA	NA	1	6	NA	2	#FOSIT
23	406	19	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
29	544	23	541	NA	NA	1	3	NA	2	#Neer et al. 2001
121	3906	101	3902	NA	NA	1	4	NA	2	#FREEDOM
14	333	5	327	NA	NA	1	5	NA	2	#CHAO
149	3861	126	3877	NA	NA	1	5	NA	2	#HORIZON
60	5057	70	5044	NA	NA	1	7	NA	2	#RUTH
21	680	15	680	NA	NA	2	3	NA	2	#VERO
37	1383	29	1353	NA	NA	1	8	NA	2	#FRAME (EU-LABEL)
91	1920	67	1923	NA	NA	6	8	NA	2	#ARCH (EU-LABEL)

Non-vertebral fractures 24 months

list(NT=9,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
33	821	24	818	18	824	1	3	9	3	#ACTIVE
38	680	25	680	NA	NA	2	3	NA	2	#VERO
46	406	29	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
29	350	28	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
53	544	34	541	NA	NA	1	3	NA	2	#Neer et al. 2001
227	3906	179	3902	NA	NA	1	4	NA	2	#FREEDOM
281	3861	221	3875	NA	NA	1	5	NA	2	#HORIZON
150	5057	142	5044	NA	NA	1	7	NA	2	#RUTH
37	958	19	950	NA	NA	1	6	NA	2	#FOSIT
151	1920	124	1923	NA	NA	6	8	NA	2	#ARCH (EU-LABEL)

Non-vertebral fractures 36 months

list(NT=7,NS=9)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
51	406	36	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)
52	815	33	812	NA	NA	1	2	NA	2	#VERT-MN (North America)
293	3906	238	3902	NA	NA	1	4	NA	2	#FREEDOM
388	3861	292	3875	NA	NA	1	5	NA	2	#HORIZON
442	3223	283	3236	NA	NA	1	3	NA	2	#FIT I+II
48	333	28	327	NA	NA	1	5	NA	2	#Chao et al. 2013
214	5057	213	5044	NA	NA	1	6	NA	2	#RUTH
119	1885	109	1849	NA	NA	1	6	NA	2	#Silverman et al. 2008
207	1920	171	1923	NA	NA	3	7	NA	2	#ARCH (EU-LABEL)

Hip fractures 12 months

list(NT=6,NS=7)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
23	3906	12	3902	NA	NA	1	2	NA	2	#FREEDOM
31	3861	24	3875	NA	NA	1	3	NA	2	#HORIZON-PFT
4	333	3	327	NA	NA	1	3	NA	2	#Chao et al. 2013
8	5057	13	5044	NA	NA	1	4	NA	2	#RUTH
22	1920	14	1923	NA	NA	5	6	NA	2	#ARCH (EU)
9	1383	3	1353	NA	NA	1	6	NA	2	#FRAME (EU)
2	707	1	716	NA	NA	1	5	NA	2	#EVA

Hip fractures 24 months

list(NT=9,NS=10)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
2	350	5	360	NA	NA	2	3	NA	2	#Hadji et al. 2012
5	680	2	680	NA	NA	2	3	NA	2	#VERO
4	544	2	541	NA	NA	1	3	NA	2	#Neer et al. 2001
0.5	819	2.5	822	0.5	825	3	1	9	3	#ACTIVE
35	3906	16	3902	NA	NA	1	4	NA	2	#FREEDOM
21	241	12	242	NA	NA	1	5	NA	2	#Bai et al. 2013
56	3861	40	3875	NA	NA	1	5	NA	2	#HORIZON-PFT
23	5057	21	5044	NA	NA	1	7	NA	2	#RUTH
22	1005	11	1022	NA	NA	1	6	NA	2	#FIT I
42	1920	30	1923	NA	NA	6	8	NA	2	#ARCH (EU-LABEL)

Hip fractures 36 months

list(NT=7,NS=8)

r[,1]	n[,1]	r[,2]	n[,2]	r[,3]	n[,3]	t[,1]	t[,2]	t[,3]	na[]	
92	3906	26	3902	NA	NA	1	4	NA	2	#FREEDOM
88	3861	52	3875	NA	NA	1	5	NA	2	#HORIZON
22	1005	11	1022	NA	NA	1	3	NA	2	#FIT I
64	1920	40	1923	NA	NA	3	7	NA	2	#ARCH (EU-LABEL)
11	406	9	406	NA	NA	1	2	NA	2	#VERT-MN (Europe)

15	815	12	812	NA	NA	1	2	NA	2	#VERT-MN (North America)
13	333	8	327	NA	NA	1	5	NA	2	#CHAO
37	5057	35	5044	NA	NA	1	6	NA	2	#RUTH

BMD input data

Total hip

Arm-level data

list(ns.a=24, ns.t=6, nt=10)

t.a[,1]	t.a[,2]	t.a[,3]	t.a[,4]	y.a[,1]	y.a[,2]	y.a[,3]	y.a[,4]	se.a[,1]	se.a[,2]	se.a[,3]	se.a[,4]	na.a[]	#study
1	3	NA	NA	-1	2.6	NA	NA	0.284	0.369	NA	NA	2	#Neer
1	4	NA	NA	-1.1	3.4	NA	NA	0.179	0.179	NA	NA	2	#DEFEND
1	5	NA	NA	0.24	1.68	NA	NA	0.43	0.41	NA	NA	2	#SPIMOS
1	5	NA	NA	-0.93	1.49	NA	NA	0.227	0.268	NA	NA	2	#McClung2009
1	6	NA	NA	-1.45	2.28	NA	NA	0.202	0.216	NA	NA	2	#NCT00132808
1	7	NA	NA	0.1	3.1	NA	NA	0.129	0.152	NA	NA	2	#FOSIT
1	7	NA	NA	-3.12	3.69	NA	NA	1.24	1.12	NA	NA	2	#Adami1995
1	7	NA	NA	-0.86	4.97	NA	NA	0.42	0.36	NA	NA	2	#Tucci
1	8	NA	NA	-0.83	0.9	NA	NA	0.12	0.12	NA	NA	2	#Silverman
1	9	NA	NA	3.1	8.2	NA	NA	0.1	0.153	NA	NA	2	#FRAME
2	3	NA	NA	2.05	0.83	NA	NA	0.4	0.5	NA	NA	2	#Hadji2012
3	4	NA	NA	2	3.2	NA	NA	0.549	0.539	NA	NA	2	#DATA
3	9	NA	NA	-0.4	2.9	NA	NA	0.255	0.255	NA	NA	2	#STRUCTURE

4	5	NA	NA	2.2	0.9	NA	NA	0.128	0.128	NA	NA	2 #Recknor
4	6	NA	NA	1.9	0.6	NA	NA	0.128	0.128	NA	NA	2 #Miller2016
4	7	NA	NA	3.5	2.6	NA	NA	0.102	0.102	NA	NA	2 #DECIDE
4	7	NA	NA	1.9	1.05	NA	NA	0.145	0.148	NA	NA	2 #STAND
6	7	NA	NA	20.1	9.1	NA	NA	1.179	0.467	NA	NA	2 #Tan
7	9	NA	NA	3.4	7.3	NA	NA	0.102	0.102	NA	NA	2 #ARCH
1	3	8	NA	0.3	1.9	1.5	NA	0.434	0.255	0.459	NA	3 #EUROFORS
1	4	7	NA	-3.52	6.06	1.17	NA	0.65	0.56	0.62	NA	3 #Amgen20010223
1	2	7	NA	-0.17	0.93	2.7	NA	0.344	0.253	0.263	NA	3 #FACTS1
1	3	10	NA	0.07	2.81	3.45	NA	0.03	0.12	0.12	NA	3 #ACTIVE
1	3	7	9	-0.7	1.3	1.9	4.1	0.332	0.332	0.332	0.332	4 #McClung2014

Trial-level data

t[,1]	t[,2]	t[,3]	y[,2]	y[,3]	se[,2]	se[,3]	na[]	#study
1	6	NA	6.02	NA	0.13	NA	2	#HORIZON
1	6	NA	3.5	NA	0.689	NA	2	#Grey
2	4	NA	1.6	NA	0.204	NA	2	#Roux2013
2	7	NA	1.7	NA	0.23	NA	2	#FACT
5	7	NA	-0.14	NA	0.135	NA	2	#MOTION
8	7	NA	1.6	NA	0.332	NA	2	#EFFECTinternational

Femoral Neck

Arm-level data

list(ns.a=26, ns.t=5, nt=10)

t.a[,1]	t.a[,2]	t.a[,3]	t.a[,4]	y.a[,1]	y.a[,2]	y.a[,3]	y.a[,4]	se.a[,1]	se.a[,2]	se.a[,3]	se.a[,4]	na.a[]
1	2	NA	NA	-1	1.3	NA	NA	0.32	0.33	NA	NA	2
#Fogelman												
1	3	NA	NA	-0.7	2.8	NA	NA	0.247	0.26	NA	NA	2
#Neer												
1	4	NA	NA	-0.9	2.8	NA	NA	0.281	0.255	NA	NA	2
#DEFEND												
1	5	NA	NA	-0.75	1.09	NA	NA	0.457	0.327	NA	NA	2
#McClung2009												
1	6	NA	NA	-1.35	1.64	NA	NA	0.288	0.308	NA	NA	2
#NCT00132808												
1	7	NA	NA	2.33	3.75	NA	NA	0.611	0.863	NA	NA	2
#Dursun												
1	7	NA	NA	-0.2	2.3	NA	NA	0.151	0.153	NA	NA	2
#FOSIT												
1	7	NA	NA	-2.58	1.19	NA	NA	0.89	0.88	NA	NA	2
#Adami1995												
1	7	NA	NA	-0.56	2.92	NA	NA	0.22	0.23	NA	NA	2
#Aki												
1	7	NA	NA	-1.6	4.66	NA	NA	0.4	0.71	NA	NA	2
#Tucci												
1	7	NA	NA	-0.6	2.5	NA	NA	0.5	0.4	NA	NA	2
#NCT00398606												
1	8	NA	NA	0.2	2.3	NA	NA	0.3	0.4	NA	NA	2
#Adami2008												
1	9	NA	NA	2.2	6.7	NA	NA	0.255	0.204	NA	NA	2
#FRAME												
2	3	NA	NA	2.11	0.77	NA	NA	0.4	0.4	NA	NA	2
#Hadji2012												
3	4	NA	NA	2.8	4.1	NA	NA	0.647	0.661	NA	NA	2
#DATA												
3	9	NA	NA	-0.4	3	NA	NA	0.306	0.306	NA	NA	2
#STRUCTURE												
4	5	NA	NA	1.7	0.5	NA	NA	0.204	0.204	NA	NA	2
#Recknor												
4	7	NA	NA	2.4	1.8	NA	NA	0.128	0.153	NA	NA	2
#DECIDE												
6	7	NA	NA	13.5	6.3	NA	NA	0.942	0.316	NA	NA	2
#Tan												
7	9	NA	NA	2.3	6.1	NA	NA	0.204	0.153	NA	NA	2
#ARCH												

1	3	8	NA	1.1	2.6	1.6	NA	0.536	0.332	0.536	NA	3 #EUROFORS
1	7	8	NA	-1.6	3.1	1.9	NA	0.5	0.3	0.5	NA	3 #Um2017
1	7	8	NA	0.2	2.7	1.7	NA	0.4	0.5	0.4	NA	3 #Johnell
1	2	7	NA	-0.08	1.44	2.23	NA	0.393	0.288	0.296	NA	3 #FACTS1
1	3	10	NA	-0.44	2.24	2.9	NA	0.13	0.13	0.14	NA	3 #ACTIVE
1	3	7	9	-1.1	1.1	1.2	3.7	0.459	0.459	0.459	0.459	4 #McClung2014

Trial-level data

t[,1]	t[,2]	t[,3]	y[,2]	y[,3]	se[,2]	se[,3]	na[]	#study
1	6	NA	5.06	NA	0.153	NA	2	#HORIZON
1	7	NA	5.9	NA	0.5	NA	2	#Liberman
2	4	NA	1.4	NA	0.23	NA	2	#Roux2013
2	7	NA	1.9	NA	0.332	NA	2	#FACT
8	7	NA	1.3	NA	0.408	NA	2	#EFFECTinternational

Lumbar Spine

Arm-level data

list(ns.a=32, ns.t=9, nt=10)

t.a[,1]	t.a[,2]	t.a[,3]	t.a[,4]	y.a[,1]	y.a[,2]	y.a[,3]	y.a[,4]	se.a[,1]	se.a[,2]	se.a[,3]	se.a[,4]	na.a[] #study
1	2	NA	NA	0.05	4.49	NA	NA	0.54	0.38	NA	NA	2 #NCT00353080
1	2	NA	NA	0	4.1	NA	NA	0.35	0.35	NA	NA	2 #Fogelman
1	3	NA	NA	1.1	9.7	NA	NA	0.245	0.332	NA	NA	2 #Neer
1	4	NA	NA	-0.6	6.5	NA	NA	0.332	0.357	NA	NA	2 #DEFEND
1	5	NA	NA	0.17	4.33	NA	NA	0.54	0.58	NA	NA	2 #SPIMOS

1	6	NA	NA	-1.32	4.42	NA	NA	0.268	0.281	NA	NA	2
#NCT00132808												
1	7	NA	NA	-0.36	7.19	NA	NA	0.957	1.064	NA	NA	2
#Dursun												
1	7	NA	NA	0.1	5	NA	NA	0.113	0.108	NA	NA	2
#FOSIT												
1	7	NA	NA	-0.01	5.2	NA	NA	0.67	0.57	NA	NA	2
#Adami1995												
1	7	NA	NA	-1.56	4.82	NA	NA	0.52	0.99	NA	NA	2
#Aki												
1	7	NA	NA	-0.76	9.59	NA	NA	0.26	0.43	NA	NA	2
#Tucci												
1	7	NA	NA	0.8	6	NA	NA	0.4	0.3	NA	NA	2
#NCT00398606												
1	7	NA	NA	0.6	5.2	NA	NA	0.48	0.48	NA	NA	2
#OCEAN												
1	8	NA	NA	-4	-1	NA	NA	0.3	0.3	NA	NA	2
#Adami2008												
1	8	NA	NA	0.88	2.96	NA	NA	0.16	0.16	NA	NA	2
#Silverman												
1	9	NA	NA	5.3	16.4	NA	NA	0.153	0.255	NA	NA	2
#FRAME												
2	3	NA	NA	7.8	2.63	NA	NA	0.5	0.5	NA	NA	2
#Hadji2012												
3	4	NA	NA	9.5	8.3	NA	NA	1.06	0.592	NA	NA	2
#DATA												
3	9	NA	NA	5.3	9.7	NA	NA	0.357	0.357	NA	NA	2
#STRUCTURE												
4	5	NA	NA	4.1	2.1	NA	NA	0.204	0.204	NA	NA	2
#Recknor												
4	6	NA	NA	3.2	1.1	NA	NA	0.204	0.204	NA	NA	2
#Miller2016												
4	7	NA	NA	5.3	4.2	NA	NA	0.153	0.153	NA	NA	2
#DECIDE												
4	7	NA	NA	3.03	1.85	NA	NA	0.207	0.209	NA	NA	2
#STAND												
6	7	NA	NA	41.3	16.9	NA	NA	1.345	0.879	NA	NA	2
#Tan												
7	9	NA	NA	7.2	15.3	NA	NA	0.153	0.204	NA	NA	2
#ARCH												
1	3	8	NA	-2.7	-0.2	1.6	NA	0.459	0.255	0.459	NA	3
#EUROFORS												

1	4	7	NA	-2.39	10.34	4.54	NA	1.11	0.96	1.606	NA	3 #Amgen20010223
1	7	8	NA	-1.81	6.7	4.36	NA	0.2	0.5	0.3	NA	3 #Um2017
1	7	8	NA	-0.004	4.3	2.1	NA	0.3	0.4	0.4	NA	3 #Johnell
1	2	7	NA	0.09	2.8	4.75	NA	0.365	0.268	0.278	NA	3 #FACTS1
1	3	10	NA	0.5	9.1	9.22	NA	0.13	0.23	0.26	NA	3 #ACTIVE
1	3	7	9	-0.1	7.1	4.1	11.3	0.536	0.536	0.536	0.536	4 #McClung2014

Trial-level data

t[,1]	t[,2]	t[,3]	y[,2]	y[,3]	se[,2]	se[,3]	na[]	#study
1	6	NA	6.71	NA	0.523	NA	2	#HORIZON
1	6	NA	5.7	NA	0.689	NA	2	#Grey
1	6	NA	3.6	NA	0.6663	NA	2	#Reid
1	7	NA	8.8	NA	0.4	NA	2	#Liberman
2	4	NA	2.3	NA	0.255	NA	2	#Roux
2	7	NA	1.8	NA	0.332	NA	2	#FACT
5	7	NA	-0.69	NA	0.217	NA	2	#MOTION
8	7	NA	2.5	NA	0.332	NA	2	#EFFECT
8	7	NA	2.6	NA	0.459	NA	2	#EFFECTinternational

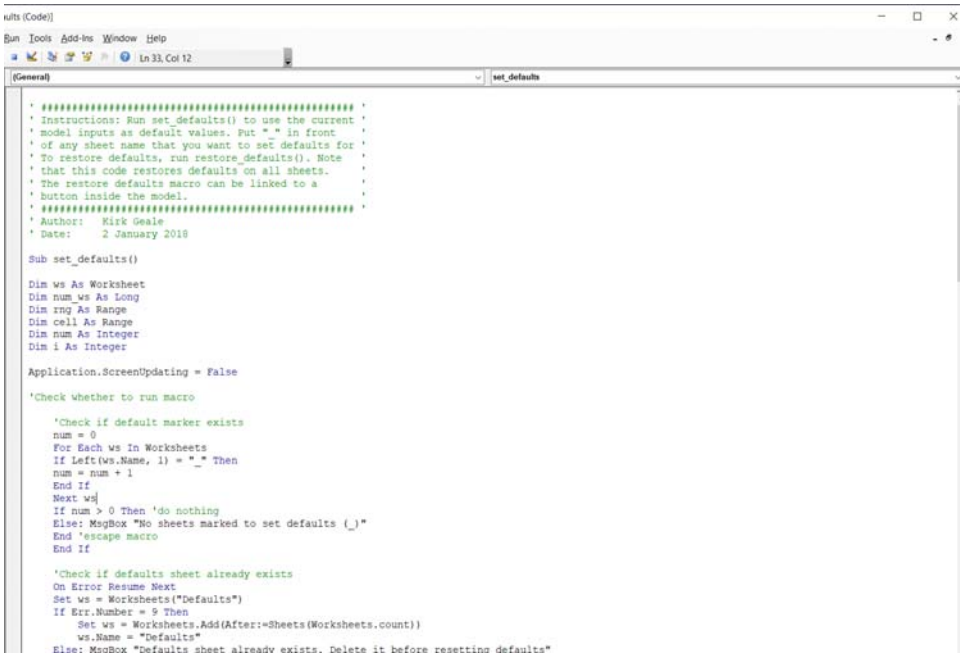
Appendix C2B

1. Module mDefaults

a. Macro set_defaults: Creates a new sheet named “Defaults” that contain cells values and formulas from all sheets in the model. Requires that a “_” is put in front of each sheet that the user wants to save default values/formulas from.

b. Macro restore_defaults: Restore default values from sheet “Defaults” to the relevant sheets.

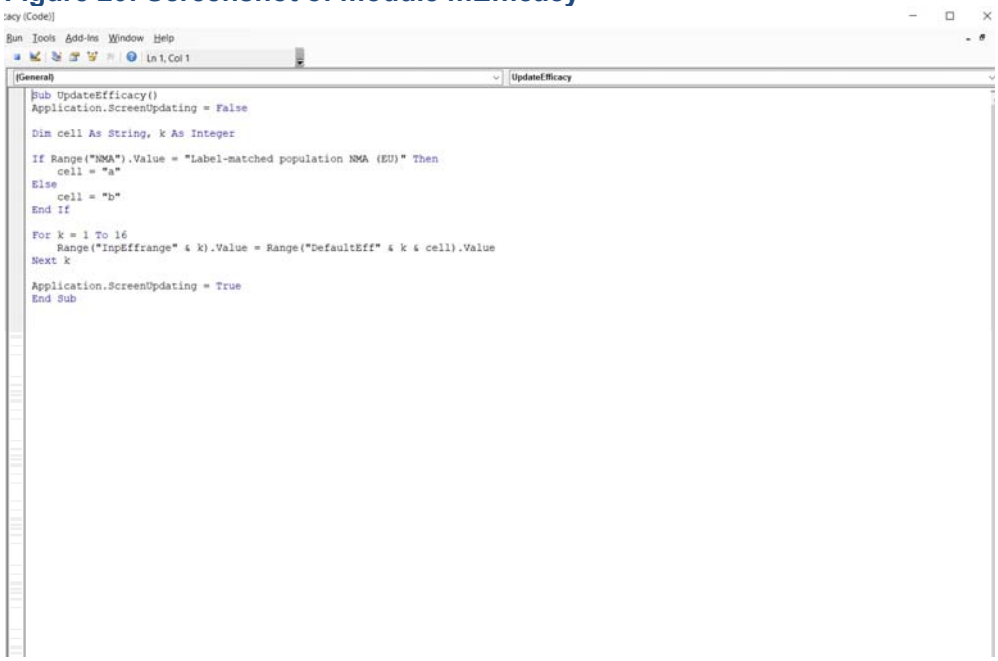
Figure 19: Screenshot of module mDefaults



2. Module mEfficacy

a. Macro UpdateEfficacy: Updates the efficacy input on sheet “Efficacy input” when NMA is switched (ITT or label-matched population), using the drop-down menu on the same sheet. The button “Update efficacy input” on the same sheet is linked to this macro.

Figure 20: Screenshot of module mEfficacy



3. Module MiscFunctions

a. Macro ShowDialog: Initiates the dialog form that shows up when running the model

b. Macro FraxGenerator: Not relevant for the submission

c. Macro SaveSequenceProfile: Saves the treatment sequences created on sheet “Treatment sequences”. A short name for each treatment sequence is saved on the same page, which can then be chosen from the drop-down list “Treatment sequence profile” on sheet “Main settings”. The macro is called when the button “Save sequence profile” on sheet “Treatment sequences”.

d. Macro ClearSequenceProfile: Removes all saved treatment sequences, i.e., clears the table “Saved treatment sequences” on sheet “Treatment sequences”

e. Macro CheckBox3_Click: Sets cell D14 in Misc tab to "True" if the "no treatment" comparator is enabled on the Main settings tab.

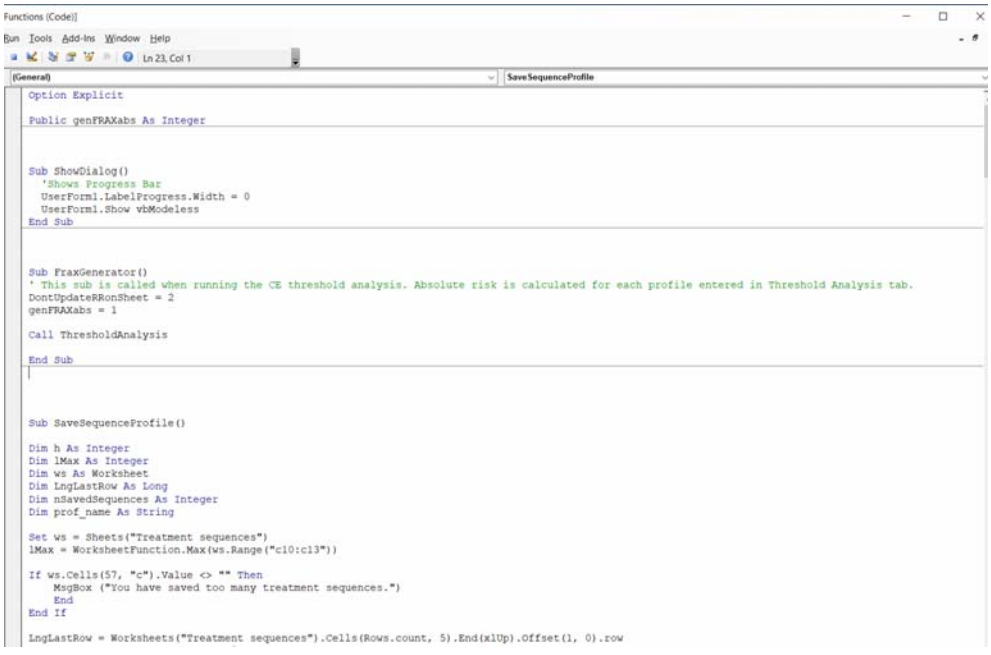
f. Macro clean: Clears all results from previous model runs. Called upon when starting a model run.

g. Macro Copy_Results: Copies the sheet “Results” to a separate workbook, when the button “Export results sheet” on sheet “Results” is pressed.

h. Macro CrossValue: Updates the base case ICER/incremental QALY/incremental costs in the graphs on sheet “DSA results”. Called upon when pressing the graphs.

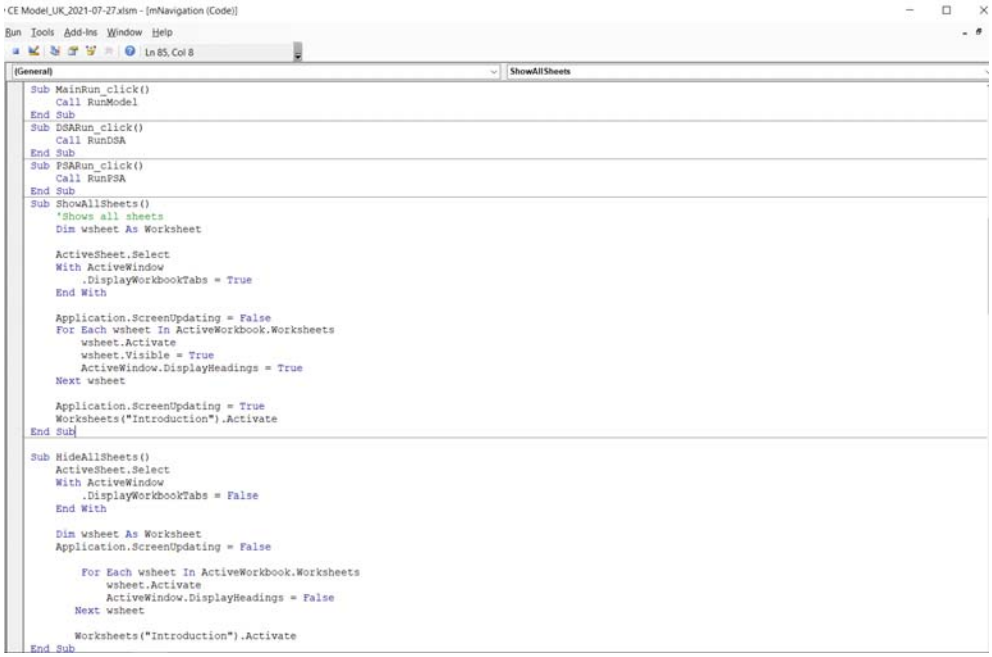
i. Macro SortDSAresults: Sorts the case ICERs/incremental QALYs/incremental costs for the deterministic sensitivity analyses n ascending order. Called upon when pressing the graphs.

Figure 21: Screenshot of module MiscFunctions



4. Module mNavigation: Contains several macros that are used when navigating the model using the buttons in the top-left corner of each sheet.

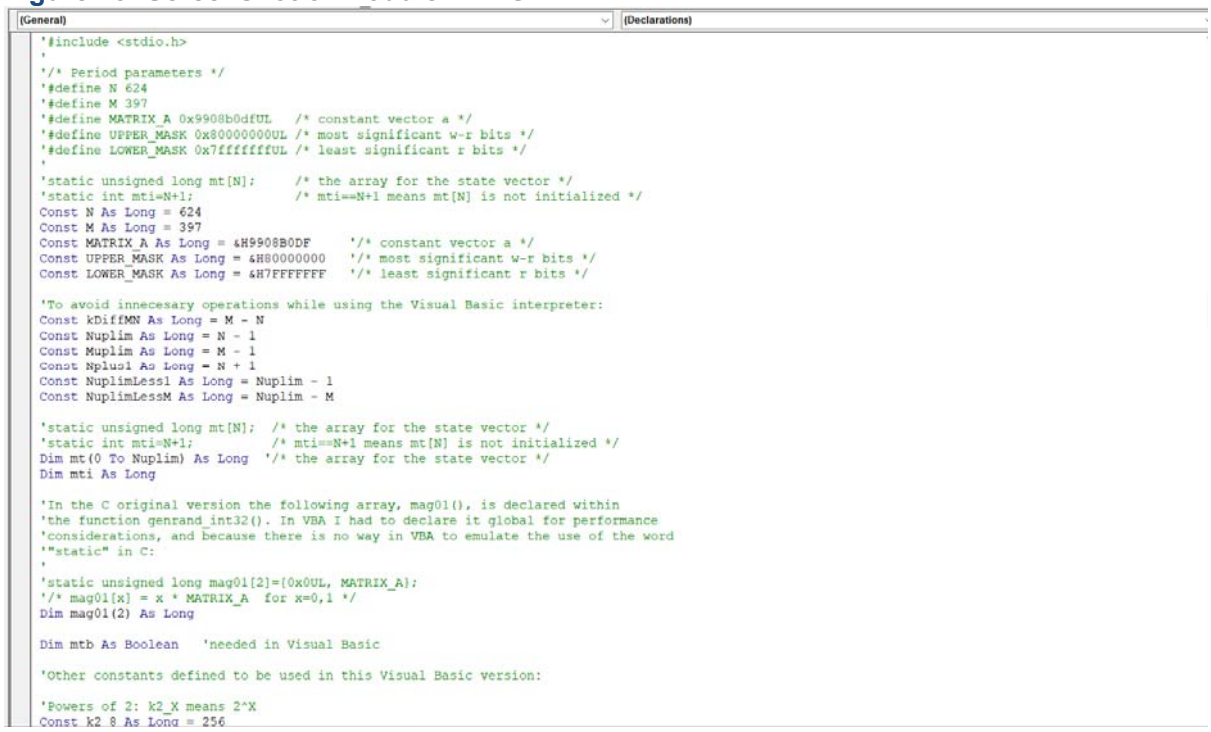
Figure 22: Screenshot of module mNavigation



```
Sub MainRun_click()
    Call RunModel
End Sub
Sub DSARun_click()
    Call RunDSA
End Sub
Sub PSARun_click()
    Call RunPSA
End Sub
Sub ShowAllSheets()
    'Shows all sheets
    Dim wsheet As Worksheet

    ActiveSheet.Select
    With ActiveWindow
        .DisplayWorkbookTabs = True
    End With

    Application.ScreenUpdating = False
    For Each wsheet In ActiveWorkbook.Worksheets
        wsheet.Activate
        wsheet.Visible = True
        ActiveWindow.DisplayHeadings = True
    Next wsheet

    Application.ScreenUpdating = True
    Worksheets("Introduction").Activate
End Sub

Sub HideAllSheets()
    ActiveSheet.Select
    With ActiveWindow
        .DisplayWorkbookTabs = False
    End With

    Dim wsheet As Worksheet
    Application.ScreenUpdating = False

        For Each wsheet In ActiveWorkbook.Worksheets
            wsheet.Activate
            ActiveWindow.DisplayHeadings = False
       Next wsheet

       Worksheets("Introduction").Activate
End Sub
```

5. Module PRNG: Contains several functions for pseudo generated random numbers based on the Marsienne Twister algorithm

Figure 23: Screenshot of module PRNG



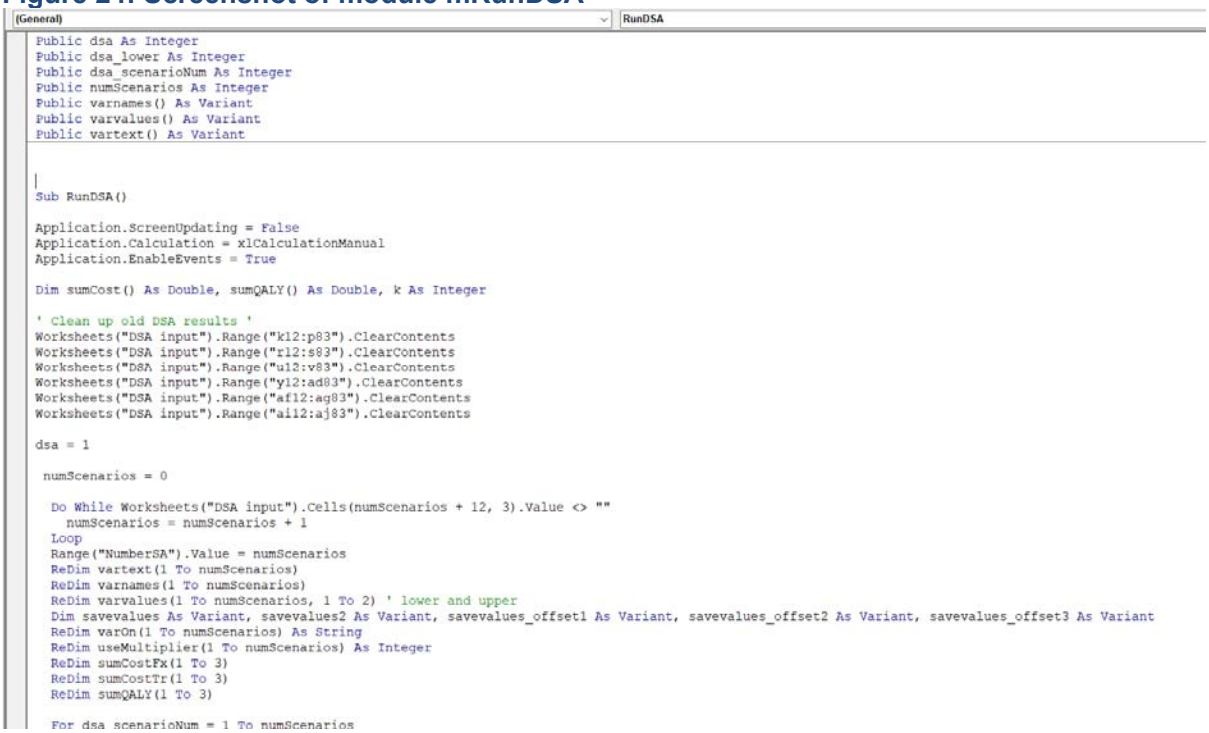
```
'#include <stdio.h>
'
'/* Period parameters */
'#define N 624
'#define M 397
'#define MATRIX_A 0x9908b0dfUL   /* constant vector a */
'#define UPPER_MASK 0x80000000UL /* most significant w-r bits */
'#define LOWER_MASK 0x7fffffffUL /* least significant r bits */
'
'static unsigned long mt[N];     /* the array for the state vector */
'static int mti=N+1;             /* mti==N+1 means mt[N] is not initialized */
Const N As Long = 624
Const M As Long = 397
Const MATRIX_A As Long = &H9908B0DF     '/* constant vector a */
Const UPPER_MASK As Long = &H80000000   '/* most significant w-r bits */
Const LOWER_MASK As Long = &H7FFFFFFF   '/* least significant r bits */

'To avoid innecesary operations while using the Visual Basic interpreter:
Const kDiffMN As Long = M - N
Const Nuplim As Long = N - 1
Const Muplim As Long = M - 1
Const Nplus1 As Long = N + 1
Const NuplimLess1 As Long = Nuplim - 1
Const NuplimLessM As Long = Nuplim - M

'static unsigned long mt[N];  /* the array for the state vector */
'static int mti=N+1;          /* mti==N+1 means mt[N] is not initialized */
Dim mt(0 To Nuplim) As Long  '/* the array for the state vector */
Dim mti As Long

'In the C original version the following array, mag01(), is declared within
'the function genrand_int32(). In VBA I had to declare it global for performance
'considerations, and because there is no way in VBA to emulate the use of the word
'"static" in C:
'
'static unsigned long mag01[2]={0x0UL, MATRIX_A};
'/* mag01[x] = x * MATRIX_A  for x=0,1 */
Dim mag01(2) As Long

Dim mtb As Boolean   'needed in Visual Basic

'Other constants defined to be used in this Visual Basic version:

'Powers of 2: k2_X means 2^X
Const k2_8 As Long = 256
```

6. Module mRunDSA

a. Macro RunDSA: Pulls in the input values used in the DSA (based on the input on sheet “DSA input”) and replaces the base case values on the other input sheets. The

base case values are temporarily stored in the memory. The main model run (macro RunModel) is than called, and the results from the DSA are stored on the “DSA input” sheet. When all DSAs have been run, the macro restores the input values to the base case.

Figure 24: Screenshot of module mRunDSA



```
(General)                                                    RunDSA

Public dsa As Integer
Public dsa_lower As Integer
Public dsa_scenarioNum As Integer
Public numScenarios As Integer
Public varnames() As Variant
Public varvalues() As Variant
Public vartext() As Variant


Sub RunDSA()

Application.ScreenUpdating = False
Application.Calculation = xlCalculationManual
Application.EnableEvents = True

Dim sumCost() As Double, sumQALY() As Double, k As Integer

' Clean up old DSA results '
Worksheets("DSA input").Range("k12:p83").ClearContents
Worksheets("DSA input").Range("r12:s83").ClearContents
Worksheets("DSA input").Range("u12:v83").ClearContents
Worksheets("DSA input").Range("y12:ad83").ClearContents
Worksheets("DSA input").Range("af12:ag83").ClearContents
Worksheets("DSA input").Range("al12:aj83").ClearContents

dsa = 1

 numScenarios = 0

  Do While Worksheets("DSA input").Cells(numScenarios + 12, 3).Value <> ""
    numScenarios = numScenarios + 1
  Loop
  Range("NumberSA").Value = numScenarios
  ReDim vartext(1 To numScenarios)
  ReDim varnames(1 To numScenarios)
  ReDim varvalues(1 To numScenarios, 1 To 2) ' lower and upper
  Dim savevalues As Variant, savevalues2 As Variant, savevalues_offset1 As Variant, savevalues_offset2 As Variant, savevalues_offset3 As Variant
  ReDim varOn(1 To numScenarios) As String
  ReDim useMultiplier(1 To numScenarios) As Integer
  ReDim sumCostFx(1 To 3)
  ReDim sumCostTr(1 To 3)
  ReDim sumQALY(1 To 3)

  For dsa_scenarioNum = 1 To numScenarios
```

7. Module mRunModel (contains the main model clockwork)

a. Macro ThresholdAnalysis is not relevant for the submission.

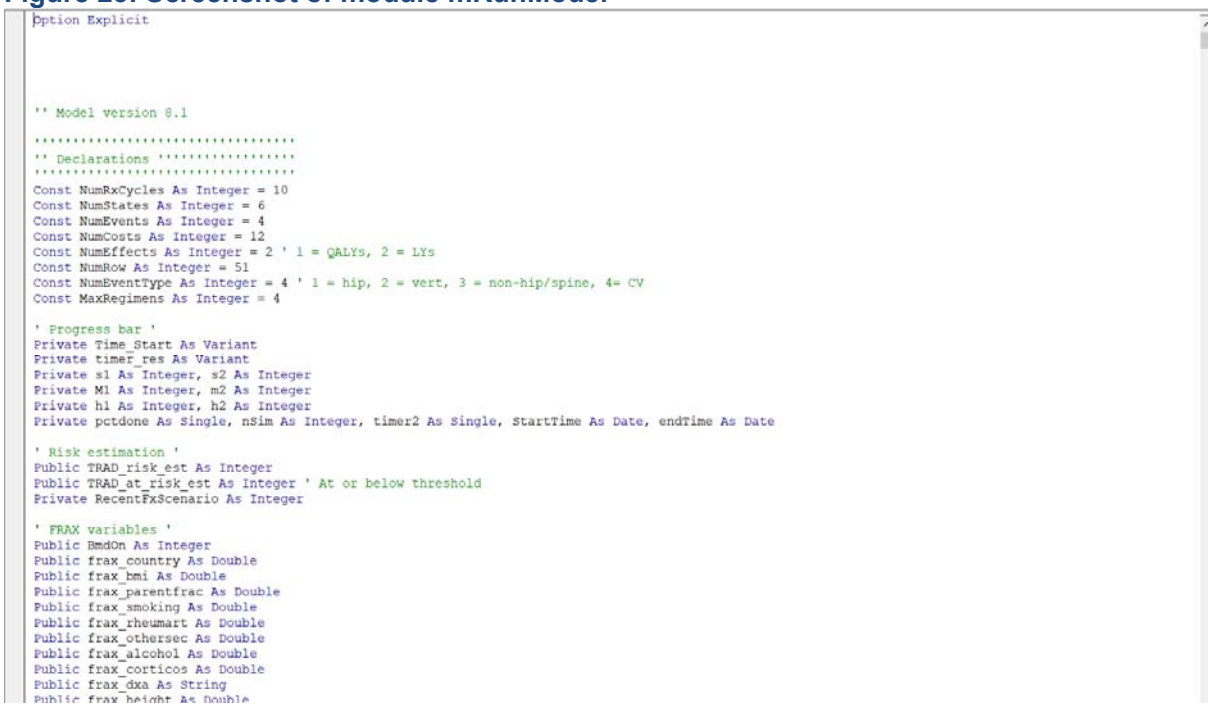
b. Macro RunModel: Initiates the model, calls macro ReadInput and TransitionPreparation, and then exports all results to the relevant sheets.

c. Macro ReadInput: Pulls all input data from the input sheets to the memory.

d. Macro TransitionPreparation: Calculates the state matrix, including randomise patients to fracture events, death. Also handles treatment specific events, such as when the patient starts and discontinues treatment, and adjusts for residual effects. Is run over three main loops: iterations (i.e., hypotethical patients), cycles (from 0 to end of time horizon), and comparators (two active treatments and no treatment).

e. Macro CalculateCostsAndEffects: Calculates costs and effects based on the state matrix calculated in macro TransitionPreparation.

Figure 25: Screenshot of module mRunModel



```
Option Explicit



'' Model version 8.1

''''''''''''''''''''''''''''''''''
'' Declarations ''''''''''''''''''
''''''''''''''''''''''''''''''''''
Const NumRxCycles As Integer = 10
Const NumStates As Integer = 6
Const NumEvents As Integer = 4
Const NumCosts As Integer = 12
Const NumEffects As Integer = 2 ' 1 = QALYs, 2 = LYs
Const NumRow As Integer = 51
Const NumEventType As Integer = 4 ' 1 = hip, 2 = vert, 3 = non-hip/spine, 4= CV
Const MaxRegimens As Integer = 4

' Progress bar '
Private Time_Start As Variant
Private timer_res As Variant
Private s1 As Integer, s2 As Integer
Private M1 As Integer, m2 As Integer
Private h1 As Integer, h2 As Integer
Private pctdone As Single, nSim As Integer, timer2 As Single, StartTime As Date, endTime As Date

' Risk estimation '
Public TRAD_risk_est As Integer
Public TRAD_at_risk_est As Integer ' At or below threshold
Private RecentFxScenario As Integer

' FRAX variables '
Public BmdOn As Integer
Public frax_country As Double
Public frax_bmi As Double
Public frax_parentfrac As Double
Public frax_smoking As Double
Public frax_rheumart As Double
Public frax_othersec As Double
Public frax_alcohol As Double
Public frax_corticos As Double
Public frax_dxa As String
Public frax_height As Double
```

8. Module mRunPSA

a. Macro RunPSA: Pulls in the input values used in the PSA (based on the input on sheet “PSA input”) and replaces the base case values on the other input sheets. The base case values are temporarily stored in the memory. The main model run (macro RunModel) is than called, and the results from the PSA are stored on the “PSA input” sheet. When all PSA iterations have been run, the macro restores the input values to the base case.

Figure 26. Screenshot of module mRunPSA

```
Public psa As Integer
Public PSAIter As Integer
Public PSAMaxIter As Integer
Const maxPSA = 106
Const numPSAcomparators = 3


Sub RunPSA()

Application.ScreenUpdating = False
Application.Calculation = xlCalculationManual
Application.EnableEvents = True

Dim sumCost() As Double, sumQALY() As Double, k As Integer, t As Integer
Dim PSAuseRomoEff, PSAuseCompEff, PSAuseUtil, PSAuseCost, PSAusePersistence, PSAuseLTC As Boolean
Dim count As Integer, maxWTP As Double, stepsize As Double, steps As Integer, comparison As Integer

ReDim vartext(1 To maxPSA)

'ReDim useMultiplier(1 To maxPSA) As Integer
ReDim savevalues(1 To 18, 1 To 51, 1 To 9)
ReDim varnames(1 To 18)
ReDim varvalues(1 To 18, 1 To 51, 1 To 3)

ReDim sumCost(1 To 3)
ReDim sumQALY(1 To 3)

' Check which parameters to be included '
PSAuseRomoEff = Worksheets("PSA input").CheckBoxPSARomoEff
PSAuseCompEff = Worksheets("PSA input").CheckBoxPSACompEff
PSAuseUtil = Worksheets("PSA input").CheckBoxPSAUtils
PSAuseCost = Worksheets("PSA input").CheckBoxPSACosts
PSAusePersistence = Worksheets("PSA input").CheckBoxPSAPersistence
PSAuseLTC = Worksheets("PSA input").CheckBoxPSALTC

PSAMaxIter = Range("InpPSAMaxIter").Value

ReDim PSAResults(1 To PSAMaxIter, 115) As Variant

' Clean up old PSA results '
Worksheets("PSA input").Range("ah7:ew1050").ClearContents
On Error GoTo handleCancel
Application.EnableCancelKey = xlErrorHandler
psa = 1
```

9. Module mUpdateRR

a. Macro updateRR: Updates the relative risks shown on sheet “Main settings” based on the inputs on the same sheet on risk profile of the patient population.

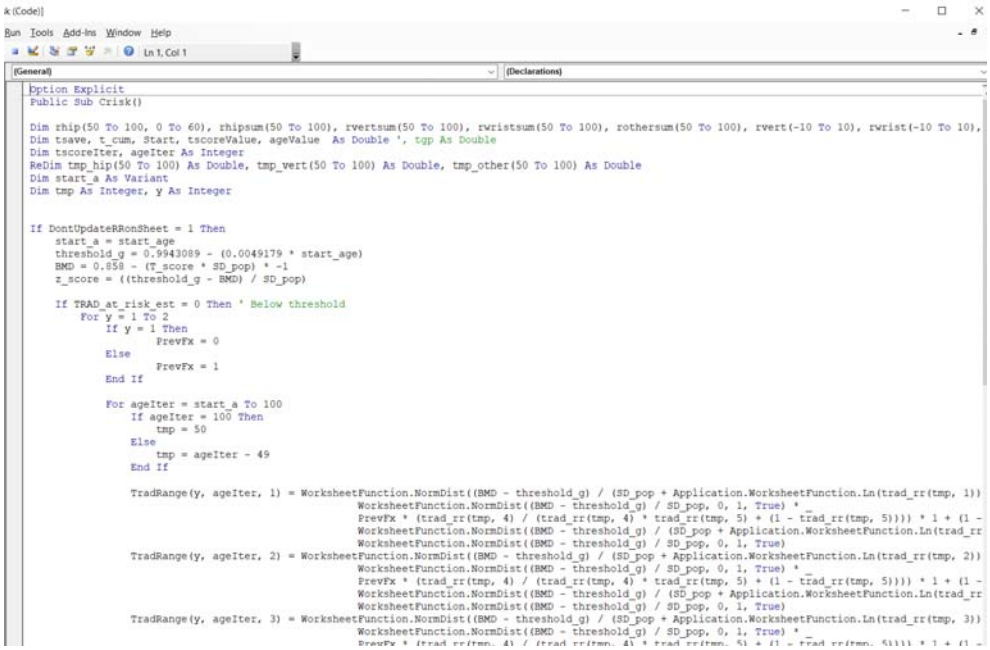
Figure 27: Screenshot of module mUpdateRR

```
Option Explicit

Sub updateRR()

Dim hip_arg1 As Double, vert_arg1 As Double, other_arg1 As Double, arg2 As Double

If DontUpdateRRonSheet <> 1 Then
    If Worksheets("Main settings").Range("risk_estimation").Value = "FRAX" Then

        Call FRAX

    Else
        MsgBox ("Only necessary when risk estimation is set to FRAX")
    End If
End If

'''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''
If DontUpdateRRonSheet = 1 Then
    If TRAD_risk_est = 0 Then ' Calculate RR
        Call FRAX

        RR_vert(comparator) = RR_major
        RR_other(comparator) = RR_major
    Else
        Call Crisk
    End If

End If

End Sub
```

10. Module: Crisk

a. Macro Crisk; Calculates the relative risks of fracture (hip, vertebral and other) when using the Traditional risk estimation method. This method is an alternative to FRAX used to calculate fracture risk in the patient population and was mainly used historically in osteoporosis models before the advent of FRAX. This method is not relevant for the submission, since FRAX is used instead of the Traditional method.

Figure 28. Screenshot of module mC_risk (macro Crisk)



```
Option Explicit
Public Sub Crisk()

Dim rhip(50 To 100, 0 To 60), rhipsum(50 To 100), rvertsum(50 To 100), rwristsum(50 To 100), rothersum(50 To 100), rvert(-10 To 10), rwrist(-10 To 10),
Dim tsave, t_cum, Start, tscoreValue, ageValue  As Double ', tgp As Double
Dim tscoreIter, ageIter As Integer
ReDim tmp_hip(50 To 100) As Double, tmp_vert(50 To 100) As Double, tmp_other(50 To 100) As Double
Dim start_a As Variant
Dim tmp As Integer, y As Integer


If DontUpdateRRonSheet = 1 Then
    start_a = start_age
    threshold_g = 0.9943089 - (0.0049179 * start_age)
    BMD = 0.858 - (T_score * SD_pop) * -1
    z_score = ((threshold_g - BMD) / SD_pop)

    If TRAD_at_risk_est = 0 Then ' Below threshold
        For y = 1 To 2
            If y = 1 Then
                    PrevFx = 0
            Else
                    PrevFx = 1
            End If

            For ageIter = start_a To 100
                If ageIter = 100 Then
                    tmp = 50
                Else
                    tmp = ageIter - 49
                End If

                TradRange(y, ageIter, 1) = WorksheetFunction.NormDist((BMD - threshold_g) / (SD_pop + Application.WorksheetFunction.Ln(trad_rr(tmp, 1))
                                           WorksheetFunction.NormDist((BMD - threshold_g) / SD_pop, 0, 1, True) * _
                                           PrevFx * (trad_rr(tmp, 4) / (trad_rr(tmp, 4) * trad_rr(tmp, 5) + (1 - trad_rr(tmp, 5)))) * 1 + (1 -
                                           WorksheetFunction.NormDist((BMD - threshold_g) / (SD_pop + Application.WorksheetFunction.Ln(trad_rr
                                           WorksheetFunction.NormDist((BMD - threshold_g) / SD_pop, 0, 1, True)
                TradRange(y, ageIter, 2) = WorksheetFunction.NormDist((BMD - threshold_g) / (SD_pop + Application.WorksheetFunction.Ln(trad_rr(tmp, 2))
                                           WorksheetFunction.NormDist((BMD - threshold_g) / SD_pop, 0, 1, True) * _
                                           PrevFx * (trad_rr(tmp, 4) / (trad_rr(tmp, 4) * trad_rr(tmp, 5) + (1 - trad_rr(tmp, 5)))) * 1 + (1 -
                                           WorksheetFunction.NormDist((BMD - threshold_g) / (SD_pop + Application.WorksheetFunction.Ln(trad_rr
                                           WorksheetFunction.NormDist((BMD - threshold_g) / SD_pop, 0, 1, True)
                TradRange(y, ageIter, 3) = WorksheetFunction.NormDist((BMD - threshold_g) / (SD_pop + Application.WorksheetFunction.Ln(trad_rr(tmp, 3))
                                           WorksheetFunction.NormDist((BMD - threshold_g) / SD_pop, 0, 1, True) * _
```

Appendix C3

Guidance on how to incorporate the comparators and the associated treatment sequences in the model is provided below.

The “Treatment sequences” sheet is used to specify the treatment sequence for romosozumab and the active comparator therapy. Up to four lines of treatment within a sequence can be specified. The user defines whether a time-point (e.g., 12, 18, 24 months) after treatment start or a fracture will trigger the sequence change and which efficacy profile should be used for each treatment within the sequence.

When a treatment sequence has been set in the input ranges, press the button “Save sequence profile” and the model will store it in the grey area below the input ranges. A profile name will be auto generated, and which can then be found in the drop-down menu in the “Main settings” sheet. The profile name chosen on the sheet “Main settings” will be used for the model simulation.

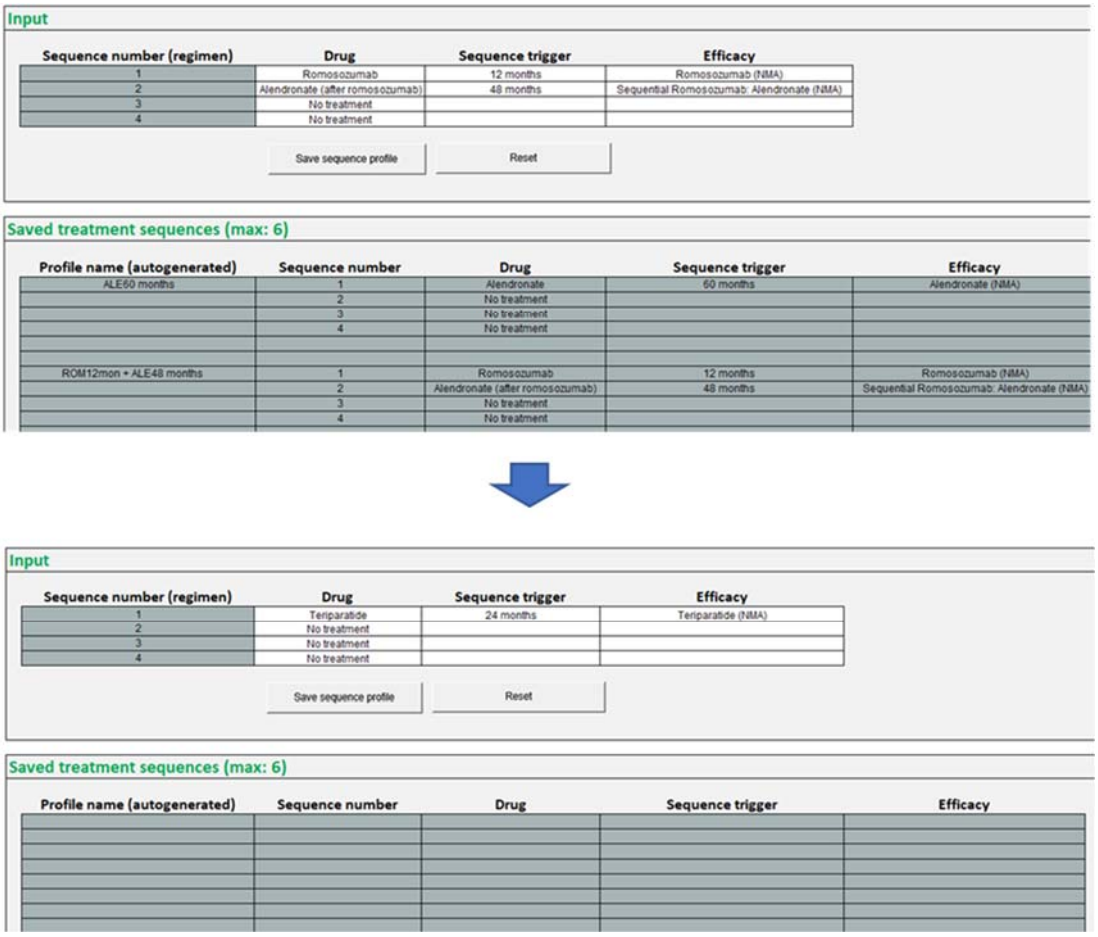
The model is pre-populated with several drug alternatives, which can be separated by being a part of a treatment sequence and being stand-alone. In the former case, the drug is named for example “Alendronate (after romosozumab)” and when alendronate is the standalone treatment, it is simply named “Alendronate”. This allows using different input profiles (e.g., efficacy, cost, persistence) depending on if the treatment is a part of a sequence or not.

Choose the relevant efficacy profile for each sequence in the drop-down menus in column F. Which efficacy profile to choose and other changes needed (from the model set for the base case) for each scenario analysis is described in the table below. An example on how to set the treatment sequence for the comparison of romosozumab-to-alendronate compared with teriparatide (24 months) is described in the screenshots below.

Step-by-step example of how to set the treatment sequence builder to the comparison of romosozumab/alendronate and teriparatide (24 months):

1. Press the Reset button to clear the table with saved treatment sequences.

Figure 29: Example screenshot of resetting the saved treatment sequences



2. Enter the sequence for teriparatide.

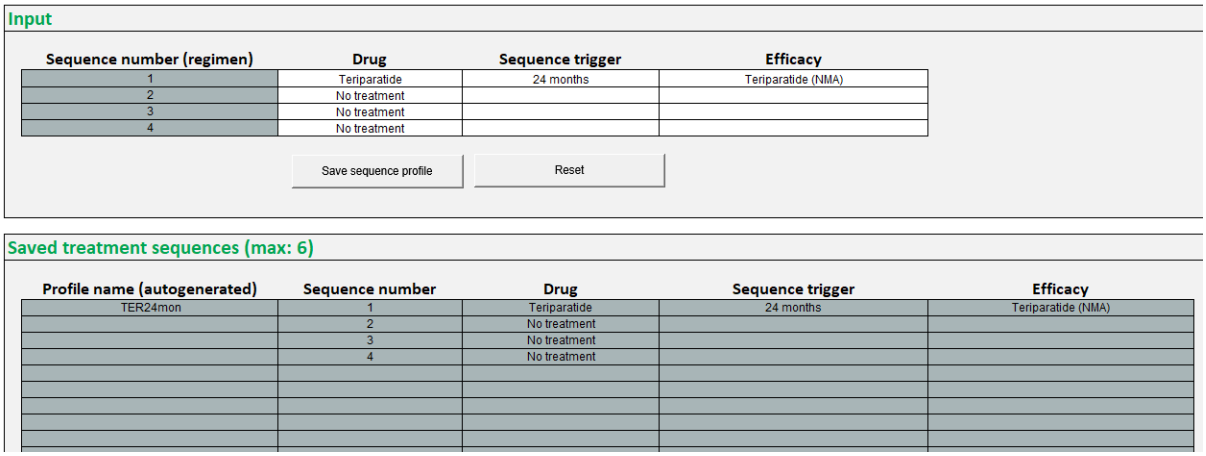
    a. Choose “Teriparatide” in the Drug list for sequence 1 and “No treatment” for sequence 2-4. Choose sequence trigger “24 months” for sequence 1 and let the cells for sequence 2-4 be empty.

    b. Choose “Teriparatide (NMA) in the Efficacy list sequence 1 and let the cells for sequence 2-4 be empty.

    c. Press Save sequence profile.

    d. The sequence shows up in the list and you can choose the profile “TER24mon” in cell E38 on Main settings.

Figure 30: Example screenshot of selecting the teriparatide sequence



Input

Sequence number (regimen)	Drug	Sequence trigger	Efficacy
1	Teriparatide	24 months	Teriparatide (NMA)
2	No treatment		
3	No treatment		
4	No treatment		

Save sequence profile | Reset

Saved treatment sequences (max: 6)

Profile name (autogenerated)	Sequence number	Drug	Sequence trigger	Efficacy
TER24mon	1	Teriparatide	24 months	Teriparatide (NMA)
	2	No treatment		
	3	No treatment		
	4	No treatment		

3. Enter the sequence for romosozumab-to-alendronate.

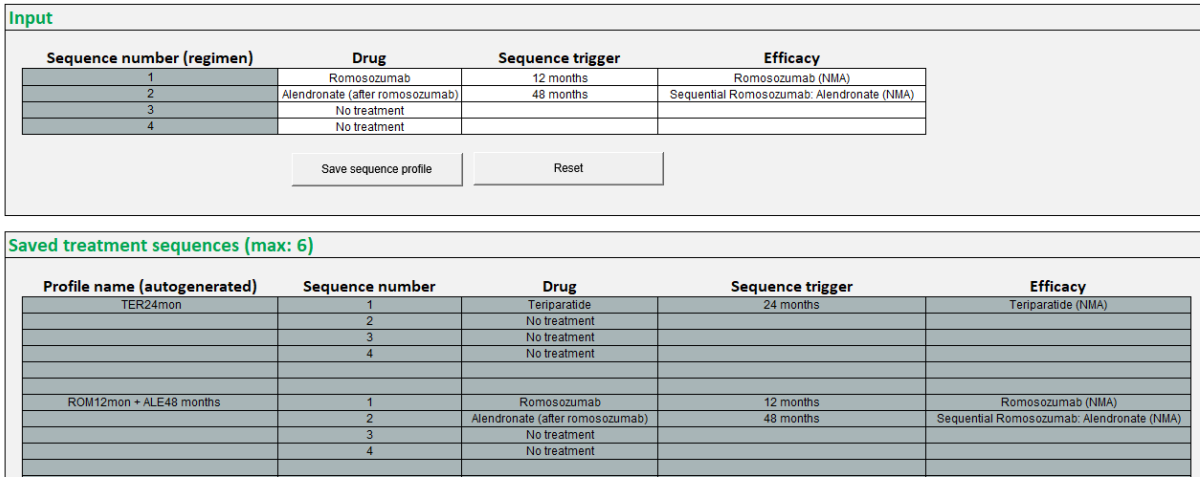
    a. Choose “Romosozumab” in the Drug list for sequence 1 and “Alendronate (after romosozumab)” for sequence 2, and “No treatment” for sequence 3-4. Choose sequence trigger “12 months” for sequence 1, “48 months” for sequence 2 and let the cells for sequence 3-4 be empty.

    b. Choose “Romosozumab (NMA)” in the Efficacy list sequence 1, “Sequential Romosozumab: Alendronate (NMA)” for sequence 2 and let the cells for sequence 3-4 be empty.

    c. Press Save sequence profile.

    d. The sequence shows up in the list and you can choose the profile “ROM12mon + ALE48 months” in cell E33 on Main settings.

Figure 31: Example screenshot of selecting the romosozumab to alendronate sequence



Input

Sequence number (regimen)	Drug	Sequence trigger	Efficacy
1	Romosozumab	12 months	Romosozumab (NMA)
2	Alendronate (after romosozumab)	48 months	Sequential Romosozumab: Alendronate (NMA)
3	No treatment		
4	No treatment		

Save sequence profile | Reset

Saved treatment sequences (max: 6)

Profile name (autogenerated)	Sequence number	Drug	Sequence trigger	Efficacy
TER24mon	1	Teriparatide	24 months	Teriparatide (NMA)
	2	No treatment		
	3	No treatment		
	4	No treatment		
ROM12mon + ALE48 months	1	Romosozumab	12 months	Romosozumab (NMA)
	2	Alendronate (after romosozumab)	48 months	Sequential Romosozumab: Alendronate (NMA)
	3	No treatment		
	4	No treatment		

Table 86: Settings by comparator and scenario analysis

Comparison	Name of efficacy profiles to choose on sheet Treatment sequences	Other settings changes required
Scenario 1: Romosozumab-to-alendronate vs. alendronate only	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Alendronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 2: Romosozumab-to-alendronate vs. Forsteo	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Teriparatide (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 3: Romosozumab-to-alendronate vs. Forsteo	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Teriparatide (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 4: Romosozumab-to-alendronate vs. raloxifene	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Raloxifene (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 5: Romosozumab-to-alendronate vs. denosumab	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Denosumab (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38. Change Efficacy offset assumption on Main settings to "Fixed" and compare the results for the denosumab arm with the results for the romosozumab arm in scenario 1.
Scenario 6: Romosozumab-to-alendronate vs. risedronate	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Risedronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.

Scenario 7: Romosozumab-to-alendronate vs. zoledronate	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Zoledronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 8: Romosozumab-to-alendronate vs. TPTD (biosimilar Movymia)-to-alendronate	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Teriparatide (Movymia, NMA) Seq. 2: Sequential Teriparatide: Alendronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 9: Romosozumab-to-alendronate vs. TPTD (Forsteo)-to-alendronate	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Teriparatide (NMA) Seq. 2: Sequential Teriparatide: Alendronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38.
Scenario 10: Romosozumab-to-alendronate vs. alendronate only	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Alendronate (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38. Switch NMA source to “Label-matched population NMA (EU)” (cell G7) and press Update efficacy input on sheet Efficacy input.
Scenario 11: Romosozumab-to-alendronate vs. denosumab	Romosozumab arm: Seq. 1: Romosozumab (NMA) Seq. 2: Sequential Romosozumab: Alendronate (NMA) Comparator arm: Seq. 1: Denosumab (NMA)	After the treatment sequence has been created, choose the appropriate sequence name in the drop-down list on Main settings E33 and E38. Change T-score on sheet Main settings I29 to -3.4. Change Efficacy offset assumption on Main settings to “Fixed” and compare the results for the denosumab arm with the results for the romosozumab arm in scenario 1.

Abbreviations: NMA: network-meta-analysis; TPTD: teriparatide.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Romosozumab for treating severe osteoporosis [ID3936]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	British Society for Rheumatology

3. Job title or position	Co-convenors of osteoporosis special interest group
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]	BSR receives funding for its biologics/biosimilars registers from Amgen, Eli Lily and Sandoz. BSR's Annual Conference 2021 (April 2021) received sponsorship funding from UCB, Amgen, Eli Lily and Novartis. BSR's Case-based Conference (October 2020) received sponsorship money from Eli Lily.

If so, please state the name of manufacturer, amount, and purpose of funding.	
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim of treatment is to strengthen bone and reduce the risk of subsequent fracture. In turn, at a population level, by lowering the number of subsequent fractures, one would expect a reduction in the mortality, morbidity and disability associated with fragility fractures, particularly those of the hip and spine.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by	Bisphosphonates, the most commonly used medications in osteoporosis, reduce major osteoporotic fracture risk by 33%, hip fracture risk by 33% and vertebral fracture risk by 55%. Risk reductions that were similar or higher than this would be clinically significant. Few drugs have been shown to reduce non-hip non-spine fractures. Any statistically significant risk reduction of non-hip non-spine fractures would be considered clinically significant.

x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. Bisphosphonates are the mainstay of treatment but are time-limited, with treatment breaks recommended after 3-10 years (depending on route of administration and severity of fracture risk). Not all patients have a satisfactory response to, or tolerate bisphosphonates, and systemic side effects (considered a class effect) are not uncommon. Denosumab is an alternative agent, but this drug is problematic due to the increased risk of rebound fracture when the drug is stopped. As a result it is now not recommended in younger people (Tsourdi *et al* JCEM 2020). Teriparatide is available as an anabolic agent, but this drug does not have proven efficacy for reducing hip fracture risk. Furthermore, access to this drug is restricted to those with low bone mineral density scores by NICE. Unfortunately, this excludes some patients with lumbar fractures who might stand to gain the most from treatment, as bone density is spuriously increased in areas of compression fracture. The unmet need can be summarised as 1. high risk patients in whom no existing drug is suitable 2. high risk patients at risk of both vertebral and hip fractures, in whom currently a combination of drugs might seem most appropriate 3. high risk patients at risk of vertebral fractures who do not meet eligibility criteria for currently available anabolic drugs.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Patients with osteoporosis, or deemed to be at high risk of fracture using available fracture risk calculators, are potentially eligible for drug treatment. National Osteoporosis Guideline Group (NOGG) guidance is used to determine thresholds for treatment. NICE guidance TA 464, TA 161 and NOGG guidance is used

	to inform treatment choice. Oral Bisphosphonates tend to be used first line, with subsequent progression to parenteral treatments, and teriparatide being used if NICE criteria are met.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	National Osteoporosis Guideline Group guidance. NICE guidance TA 464, TA 161 Scottish Intercollegiate Guideline Network Management of osteoporosis and the prevention of fragility fractures SIGN 142
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The pathway of care is variable, including referral criteria to secondary care, availability of some parenteral treatment in primary care (in a few areas denosumab can be instigated in primary care and zoledronate is given in the community) and the specialism providing specialist care (rheumatology, endocrinology, clinical biochemists, elderly care). Despite this variation, all areas have someone identified with osteoporosis special interest and a secondary care pathway for assessment for, and initiation of, secondary care prescribed parenteral treatments. There is general agreement with the principle of the need to identify higher risk individuals, particularly those with vertebral fractures, for more aggressive therapy.
• What impact would the technology have on the current pathway of care?	It would raise awareness of the need to identify higher risk individuals, particularly those with vertebral fractures, for more aggressive therapy, and may slightly increase secondary care referrals. Once referred, pathways are already established for assessing and initiating treatment.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Pathways are already established for assessing and initiating parenteral treatment, which are usually led by osteoporosis specialist nurses.

• How does healthcare resource use differ between the technology and current care?	If self-injection is not possible, the patient would need to receive subcutaneous injections monthly for 12 visits. A patient starting teriparatide, by comparison, would perhaps attend hospital on 2-3 occasions over 12 months (6 monthly appointments), and receive injections via a healthcare at home delivery service. After cessation, a sequential therapy will be needed, similar to teriparatide.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care specialist clinics, akin to use of teriparatide
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No investment in additional facilities as pathways already in place.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Romosozumab is the only drug with a dual action, stimulating both bone formation and inhibiting bone resorption. It has been demonstrated to be more effective at reducing new vertebral, non-vertebral, clinical and hip fractures than oral bisphosphonates. There is no comparison data with teriparatide, but the clinical trials have not shown teriparatide to reduce hip fracture risk, meaning romosozumab has an advantage in this regard.
• Do you expect the technology to increase length of life more than current care?	Yes, through reduced numbers of hip and spinal fractures

• Do you expect the technology to increase health-related quality of life more than current care?	Yes, through fracture reduction
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Younger people with osteoporosis (because of the need to avoid denosumab) Those unable to take bisphosphonates Those who need a drug with superior efficacy at reducing hip and vertebral fractures Those unable to take teriparatide, or previously treated with teriparatide
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	Romosozumab involves 2 subcutaneous injections administered each month for an initial period of 1 year. It is assumed that patients will be trained to self-administer these injections (pathways for training patients to self-administer injections are already in place for another anabolic osteoporosis treatment [teriparatide]). Therefore, it is anticipated that most patients would find the technology acceptable. In contrast, oral bisphosphonates (the most commonly prescribed osteoporosis medications) are often poorly tolerated, and a proportion of patients do not persist with treatment for this reason.

affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	It is anticipated that the decision to start treatment will be made on the basis of a diagnosis of severe osteoporosis based on bone density measurements (using DXA), and / or clinical characteristics such as the type or number of fragility fractures. Criteria already exist for starting teriparatide (another anabolic treatment), and it is anticipated that Romosozumab may be placed similarly. It is our understanding that a course of treatment with Romosozumab will be for 1 year. Therefore, treatment would be stopped at the end of this period, and the patient switched to an alternative (anti-resorptive) drug.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Possibly, depending on the extent to which the impact of vertebral fracture on quality of life (as opposed to the impact of hip fracture, which is well documented from a health economic perspective) is taken into account. A focus on hip fracture as an outcome may have led to previous under-valuation of osteoporosis prevention and treatment interventions.

16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	This technology will provide an alternative option for treating patients with severe osteoporosis. Currently only one other anabolic therapy for osteoporosis (teriparatide) is available in the UK, and is not suitable for all patients. The evidence for Romosozumab appears particularly strong in terms of a reduced risk of vertebral fractures; these fractures are known to be associated with significant morbidity and reduced quality of life and therefore a treatment that is effective in reducing these fractures should have a substantial clinical impact.
• Is the technology a 'step-change' in the management of the condition?	The dual effect of Romosozumab, by which it is able to simultaneously increase bone formation and reduce bone resorption (as evidenced by bone turnover marker measurements) does suggest a 'step-change' in management, as currently available anabolic therapies eventually cause an increase in bone resorption which may limit the gain in BMD achieved. However, to our knowledge there are currently no studies directly comparing these treatments in terms of fracture outcomes.
• Does the use of the technology address any particular unmet need of the patient population?	Teriparatide is not suitable for all patients, such as those with a history of prior radiotherapy to the skeleton, so this technology would offer a treatment option for these individuals. There is also a need for additional treatment options for patients at high risk of fracture who have previously received a course of teriparatide, and for patients who are unable to tolerate other treatments such as bisphosphonates.
17. How do any side effects or adverse effects of the	According to clinical studies, Romosozumab can be associated with adverse skin reactions (e.g. erythema multiforme / urticaria) and / or hypersensitivity reactions (e.g. angioedema) in some patients – this is similar

technology affect the management of the condition and the patient's quality of life?	to other treatments administered by subcutaneous injection. There is also a possible (though as yet unproven) association with cardiovascular events, which could have a significant effect on affected patients; the occurrence of such an event would require the treatment to be stopped and an alternative considered. Patients may also develop hypocalcaemia, but this should be manageable through appropriate calcium supplementation.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	There have been 2 key large-scale trials evaluating the effect of Romosozumab in the treatment of osteoporosis; FRAME and ARCH. Both trials recruited postmenopausal women, who represent the population most affected by osteoporosis in whom this treatment would be used. The FRAME study population was a comparatively lower risk group, as those with a prior hip fracture and severe / multiple vertebral fractures were excluded. The ARCH study included women with a recent hip fracture and required women to have at least one significant vertebral fracture, though BMD thresholds were less stringent. Overall, the trials did include the type of patients in whom anabolic therapy might currently be considered. However, both trials excluded women who had recently been treated with other osteoporosis therapies (e.g. oral / IV bisphosphonates) – this is a common clinical scenario in which a change in treatment would be considered, so it will be important in future studies to evaluate the effect of the treatment on fracture in this population (so far, to our knowledge, only the effect on BMD has been studied). Men were not included in the FRAME or ARCH studies, but were included in the BRIDGE phase 3 clinical trial of romosozumab *vs.* placebo, which evaluated changes in lumbar spine BMD only.

• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	The most important clinical outcomes are new hip / vertebral fractures, followed by other clinical fractures. The primary endpoint of both the FRAME and ARCH trials was new radiographic vertebral fractures, a reduction in which was observed with romosozumab treatment in both cases. The ARCH study also demonstrated a reduction in nonvertebral fractures, however the FRAME study did not. Hip fractures were commented on specifically in ARCH but not in FRAME.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	In FRAME, BMD improvement and bone turnover markers were measured in a subgroup. In ARCH, BMD was included as a secondary endpoint (and bone turnover markers measured in a subgroup). However both studies used fracture incidence as their primary outcome, which is appropriate. The BRIDGE trial of romosozumab treatment in men used lumbar spine BMD as a surrogate outcome measure. Whilst this is likely to be a good indication of future vertebral fracture risk, fracture data are needed.
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	We are not aware of any, but the treatment has yet to be used in UK clinical practice.

19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
20. How do data on real-world experience compare with the trial data?	
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	Ideally this treatment should be made available for both men and women with severe osteoporosis, however most of the clinical trials so far have only included women. In our view, this should not prevent the use of romosozumab in men, as the benefits of treatment are very likely to be similar to those in women. Previously, NICE only recommended teriparatide for women; subsequently an NHS England Clinical Commissioning policy statement in 2018 supported use in men. However, in the 10 years in between these documents, men were disadvantaged and denied access to this drug. Romosozumab has similar efficacy in East Asians to the global population (Lau et al, Osteoporosis Int 2020)

21b. Consider whether these issues are different from issues with current care and why.	A relative paucity of evidence regarding the effectiveness of osteoporosis treatments in men is common to most currently available osteoporosis therapies.

Key messages

22. In up to 5 bullet points, please summarise the key messages of your submission.

- There is a significant clinical need for a further treatment option in osteoporosis which provides an alternative to existing treatments and has proven efficacy in reducing both hip and spinal fractures
- The new technology would be expected to fit well within existing secondary cares services, and not require any additional facilities or resources.
- It is important that any economic evaluation considers the impact of reduced numbers of vertebral fractures on quality of life.
- In the issues of equality, we would hope that access for both men and women can be considered.
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Patient organisation submission

Romosozumab for treating severe osteoporosis [ID3936]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Royal Osteoporosis Society
3. Job title or position	Service Improvement Lead
4a. Brief description of the organisation (including who funds it). How many members does it have?	The Royal Osteoporosis Society (ROS) is the UK's only national charity dedicated to bone health and osteoporosis. We work to improve the bone health of the nation and support everyone with osteoporosis to live well through our support services and advice. The ROS provides both printed and digital information to help people understand more about living with osteoporosis. There is also a dedicated nurse specialist helpline and support groups locally across the UK. We influence and shape policy and practice at every level through our work with healthcare professionals and policymakers. We are driving research and development of new treatments, working towards a future without osteoporosis. We fund our work through a range of income streams – including traditional fundraising activities such as appeals and community fundraising, our membership programme, and education and training events for healthcare professionals. In a typical year, around a half of our income comes from gifts in wills, and we are extremely grateful to supporters who choose to remember us in this way. Our membership programme, individual donations, and fundraising activities such as appeals, lotteries and challenge events contribute around a third of our funding. Each year, we apply for funding to a range of national and regional charitable trusts and foundations which kindly contribute both to new projects and ongoing work. We also work with a small number of carefully selected corporate partners from the field of osteoporosis and bone health and in 2019. In 2019, we raised just over £4.2m towards our work. More detail can be found in our accounts and Trustees' Annual Report, which is available both on our website and on the Charity Commission site. A list of corporate partners can be found on our website.

	We currently have over 20,000 members
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and purpose of funding.	Yes <u>UCB - £48,295</u> £1,550 - Sponsored x3 clinical network meetings £30,070 - Grant to support Public Affairs & Advocacy programme £10,675 - Webinar series on fracture prevention services £6,000 - cost to support development of RCGP module <u>Amgen - £31,100</u> £1,100 - Sponsored x3 clinical network meetings £30,000 - Grant to support Public Affairs & Advocacy programme
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and	In 2014, the ROS (previously known as National Osteoporosis Society) conducted research into ways osteoporosis affects people's lives and published 'Life with Osteoporosis'. This report was based on the experiences of 3228 people who completed a detailed questionnaire and 52 people who took part in an in-depth interview or kept a personal diary. This survey is going to be repeated later in 2021.

carers to include in your submission?	The ROS helpline takes many calls every day from people asking questions about drug treatments, possible side effects and what alternatives there are available to their current treatment. In my role with ROS, I have facilitated meetings for people newly diagnosed with osteoporosis and had the opportunity to discuss drug treatments which are available with them. I hear issues about side effects of current medications and the need for new treatments to be available. I also hear about the impact osteoporosis and fractures have on people's daily lives. I have attended ROS Support Group meetings, and again heard people expressing concerns that they cannot tolerate their current treatment and the need they have for new alternative drug treatments. I have recently had opportunity to meet osteoporosis clinicians in Scotland who have been using Romosozumab to treat patients with severe osteoporosis for several months since its approval by the Scottish Medicines Consortium.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Osteoporosis is a condition where bones lose strength, making people affected more likely to break a bone after a minor bump or fall than the average adult. One in two women and one in five men over the age of 50 are expected to break a bone during their lifetime. Spinal fractures are the most common osteoporotic fractures; yet up to 70% of spinal fractures are not diagnosed, leaving thousands suffering, untreated and at high risk of further debilitating fractures. For many, osteoporosis means living in pain, or the fear of pain from spinal fractures. Pain, fear and fractures mean losing things in life they love. It means giving up activities, hobbies, friendships and work. People can become inactive, exacerbating the decline in their bone health. They can also struggle financially if they lose their income.

	People who experience height loss and spine curvature from spinal fractures can hate the way they look, making them feel insecure and self-conscious. A third of people with spine fractures become breathless and struggle to eat. People who suffer hip and spinal fractures have a decreased life expectancy. After a hip fracture, 7% of people die within a month and 29% within 1 year. Of those who survive, 60% can no longer carry out basic tasks for themselves such as dressing, feeding themselves and going to the toilet. A third will never return home and 43% will no longer be able to walk independently. People can feel socially isolated. Relationships can become strained as people become more dependant. Osteoporosis can stop people from seeing family & friends. For carers, it can be very difficult to watch their loved ones struggling with pain and disability as a result of fractures caused by osteoporosis. In some cases, they have to take on a new role as carer and undertake new tasks around the house and garden which may be new and unexpected to them. Patient quotes: 'Nobody understands how debilitating pain (after spine fractures) can be. I get scared and very depressed. I often cry a lot and cannot do the things a woman of my age should be able to do. I feel alone…..' 'I cannot physically hug my children and grandchildren.' 'When I found out I had fractured my spine, I had to quit my job.' 'I'm ruining my husband's life. He has to care for me when we should be enjoying our retirement.'

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Oral bisphosphonates, used as first-line treatment in the majority of individuals with osteoporosis, are sometimes not tolerated due to gastrointestinal side effects and are unsuitable in the presence of co-morbidities such as Barrett's oesophagus or significant renal impairment. These treatments must be taken on an empty stomach at least 30 minutes before food and many people find this challenging and may lead to missed doses. There are a proportion of patients who continue to fracture or experience a decline in bone mineral density (BMD) despite oral bisphosphonates. These patients then need to progress to second line treatment. Second line treatments include Zoledronic acid (administered intravenously, usually within the hospital setting, on an annual basis usually for a total of three years), so people are required to attend a clinic annually and get blood tests prior to the infusion. This can be a burden on patients who are frail and elderly who may not be able to access this treatment easily. Denosumab is another second line treatment in the form of a six-monthly subcutaneous injection usually given by a Healthcare Professional but can be self-injected. Delay or discontinuation of treatment can precipitate rapid bone loss and vertebral fractures, so patients need to remember when they are next due a dose in order to avoid inadvertent delays in treatment and the consequent risk of vertebral fracture. This has been a particular challenge causing much anxiety during the pandemic for patients unable to make timely appointments when their treatment is due. Until now, teriparatide has been the only treatment used in the management of osteoporosis with anabolic or bone-forming potential. It is administered as a daily self-administered subcutaneous injection for two years. Many patients find this challenging, especially if they are unable to manage the injections themselves (e.g., due to poor eyesight or reduced dexterity) and have to be dependent on a family member or healthcare profession to administer treatment every day. Use of teriparatide is limited to patients with very severe osteoporosis who need to fulfil several clinical criteria. In practice, teriparatide is

	generally given as a “salvage” treatment at a stage in their disease when patients are already experiencing severe pain and disability. Treatment can be very effective in preventing further deterioration but cannot reverse existing damage. This treatment is particularly helpful for those at high risk of vertebral fracture but is less useful for those who are also at high risk of hip and other non-vertebral fractures due to more modest efficacy at these sites. Patient quote: ' I took Alendronic acid for nearly 2 years, but I suffered diarrhoea on a daily basis. I didn't like to go far from home as I needed to find a toilet quickly. It wore off after I stopped taking Alendronic acid. I then changed onto Zoledronic acid and have received one infusion and didn’t feel great for a few weeks afterwards but have been fine since and am now due another one.'
8. Is there an unmet need for patients with this condition?	Yes, there are some people with osteoporosis not currently taking any treatment to reduce their risk of future fractures as they have been unable to tolerate the treatments currently available. There are also patients with very severe osteoporosis who continue to experience debilitating fractures despite complete adherence to anti-resorptive treatments. Some of these are not suitable or eligible for treatment with teriparatide or may have received this agent in the past and are unable to have a further course of treatment in accordance with the marketing authorisation.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	New treatments are always welcomed by people with osteoporosis, especially as it is the first new treatment for osteoporosis to become available for 10 years. The importance of Personalised Medicine has become recognised across many areas of medicine and osteoporosis management is no exception, requiring a range of treatment options applicable at different stages of the disease process and life course. Lived experience from Scotland has shown that people like that romosozumab is only given as a once monthly injection. This is much more acceptable than the daily injection regime of teriparatide, the only other anabolic agent. The evidence showing superiority in fracture reduction with romosozumab in

	comparison to standard osteoporosis treatment (alendronic acid) also inspires hope in those with severe disease. Last year, over 3000 people spoke to a nurse on the ROS helpline about drug treatments. Treatment gives people hope and increases their confidence to carry on activities, volunteering and work, benefiting the individual and their loved ones. Patient quote: 'Romosozumab gives another treatment option for people like me with osteoporosis and offers family and carers added confidence that quality of life for loved ones can be enhanced. It is good to know fractures can be significantly reduced and for those who are unable to tolerate other forms of osteoporosis medication this can offer renewed hope'. Patient quotes from Scotland: ‘I am managing Romosozumab well. I feel that it’s easy to administer and have not had any side effects.’ ‘I think Romosozumab is great and easy to use. I’ve not had any adverse effects.’
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	The disadvantages are that the patient and/or carer will have to be taught how to administer this subcutaneous injection where most other treatments involve taking a tablet, or the injection is administered by a Health Professional. It may be confusing for patients/carers as the treatment only lasts for one year and then they will be required to change onto a different treatment. The association between romosozumab and cardiovascular disease (CVD) adverse events is a concern to patients. Some will be deterred from using this treatment, even if their cardiovascular risk is low,

	whereas others are concerned that even minor CVD risk factors will prevent them from accessing treatment for their bone disease.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	The clinical trials indicate that the people who are likely to benefit most are those who are at the highest risk of vertebral fracture. This will include those with one or more prior vertebral fractures, particularly if these are recent and/or severe and who also have low bone mineral density. After the 12-month course of romosozumab, the trials show that the beneficial effect of treatment can be maintained by transitioning to an anti-resorptive agent. Treatment is likely to be less beneficial in the longer term if anti-resorptive agents are contraindicated or not tolerated.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Romosozumab is only licensed for post-menopausal women so cannot be used in men with osteoporosis or pre-menopausal women. We hope that future changes in the market authorisation to include these groups would be accommodated within the guidance. This treatment has already been approved by the Scottish Medicines Consortium and is now in use across Scotland. It is now also on the Formulary in Northern Ireland and has started being used there. Inability to access treatment in rest of the UK would be perceived as a “postcode lottery.

Other issues	
13. Are there any other issues that you would like the committee to consider?	No

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- Osteoporosis is a common condition which is underdiagnosed and undertreated, with many people not receiving a diagnosis until the condition is advanced.
- Hip and spinal fractures are the most serious outcomes of osteoporosis and can severely affect quality of life and decrease life expectancy.
- Treatment options for those with severe or progressive osteoporosis are limited, with only one anabolic treatment currently available to patients in England and Wales.
- The availability of Romosozumab, as the first new treatment for people with osteoporosis in 10 years, offers a step-change in the management of this debilitating condition.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Romosozumab for treating severe osteoporosis [ID3936]

Produced by	Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Rob Riemsma, Reviews Manager, KSR Ltd, United Kingdom (UK) Pim Wetzelaer, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), EUR, the Netherlands Sean Harrison, Systematic Reviewer and Statistician, KSR Ltd, UK Susan O'Meara, Systematic Reviewer, KSR Ltd, UK Hannah Penton, Health Economics Researcher, ESHPM, EUR, the Netherlands Steve Ryder, Health Economist, KSR Ltd, UK Steven Duffy, Information Specialist, KSR Ltd, UK Nigel Armstrong, Health Economics Manager, KSR Ltd, UK Maiwenn Al, Health Economics Researcher, ESHPM, EUR, the Netherlands Jos Kleijnen, Director, KSR Ltd, UK
Correspondence to	Rob Riemsma, Kleijnen Systematic Reviews Ltd Unit 6, Escrick Business Park Riccall Road, Escrick York, YO19 6FD United Kingdom
Date completed	05/10/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number STA 13/51/35.

Declared competing interests of the authors

None.

Acknowledgements

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout 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 Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Riemsma R, Wetzelaer P, Corro Ramos I, Harrison S, O'Meara S, Penton H, Ryder S, Duffy S, Armstrong N, Al M, Kleijnen J. Romosozumab for treating severe osteoporosis [ID3936]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Pim Wetzelaer acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Isaac Corro Ramos, Hannah Penton, Steve Ryder and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Sean Harrison acted as systematic reviewer and statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Susan O'Meara acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Maiwenn Al critiqued the company's economic evaluation, contributed to the writing of the report, and provided general health economic guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse event
AFF	Atypical femur fractures
AiC	Academic in confidence
ALN	Alendronate
ARR	Absolute risk reduction
ASBMR	American Society for Bone and Mineral Research
AUS	Australia
AWMSG	All Wales Medicines Strategy Group
β-CTX	Beta-C-Terminal Telopeptide of Type 1 Collagen
BC	Base-case
BMI	Body mass index
BMD	Bone mineral density
BNF	British National Formulary
BP	Bisphosphonates
BPI	Brief Pain Inventory
BTM	Bone turnover marker
CABG	Coronary artery bypass graft
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Central Register of Controlled Trials
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CG	Clinical guideline
CHD	Coronary heart disease
CHMP	Committee for Medicinal Products for Human Use
CiC	Commercial in confidence
CPRD	Clinical Practice Research data
CRD	Centre for Reviews and Dissemination
CPI	Consumer price index
Crl	Credible interval
CS	Company submission
CSR	Clinical study report
CV	Cardiovascular
DARE	Database of Abstracts of Reviews of Effects
DEN	Denosumab
DSA	Deterministic sensitivity analysis
DXA	Dual-energy x-ray absorptiometry
ECTS	European Calcified Tissue Society
EED	Economic Evaluation Database
EEPIA	European Federation of Pharmaceutical Industry Associations
EMA	European Medicines Agency
EQ-5D	EuroQoL-5 Dimensions
EQ-5D-3L	EuroQoL-5 Dimensions-3 Levels
EQ-5D-5L	EuroQoL-5 Dimensions-5 Levels
EQ-VAS	EuroQoL-Visual analogue scale
ERG	Evidence Review Group
ESCEO	European Society for Clinical and Economic Aspects of Osteoporosis
ESHPM	Erasmus School of Health Policy & Management
EU	European Union
EULAR	European League Against Rheumatism
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FFN	Fragility Fracture Network

FN	Femoral neck
FRAX	Fracture Risk Assessment tool
FSH	Follicle-stimulating hormone
GPRD	General Practice Research Database
GI	Gastrointestinal
GIAE	Gastrointestinal adverse event
GIN	Guidelines International Network
GP	General practitioner
HES	Hospital episodes statistics
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
ICUROS	International Costs and Utilities Related to Osteoporotic Fractures Study
iMTA	Institute for Medical Technology Assessment
INAHTA	International Network of Agencies for Health Technology Assessment
Inc.	Incremental
IOF	International Osteoporosis Foundation
IQR	Interquartile range
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITT	Intention-to-treat
IU	International unit
IV	Intravenous
IVRS	Interactive voice response system
KM	Kaplan Meier
KSR	Kleijnen Systematic Reviews
LAD	Limited activity days
LOCF	Last observation carried forward
LS	Lumbar spine
LTC	Long-term care
LYG	Life years gained
m	Months
MD	Mean difference
MeSH	Medical subject headings
MI	Myocardial infarction
MOF	Major osteoporotic fracture
MTA	Multiple technology appraisal
NA	Not applicable
NAm	North America
NCPE	National Centre for Pharmacoeconomics
NHNV	Non-hip, non-vertebral
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NLM	National Library of Medicine
NMA	Network meta-analysis
NOF	National Osteoporosis Foundation
NOGG	National Osteoporosis Guideline Group
NOS	National Osteoporosis Society
NR	Not reported
ONJ	Osteonecrosis of the jaw
ONS	Office for National Statistics
OPAQ-SV	Osteoporosis Assessment Questionnaire Short Version
OR	Odds ratio

OS	Overall survival
P1NP	Procollagen Type 1 N-Telopeptide
PAS	Patient access scheme
PASS	Post-Authorization Safety Studies
PICOS	Population, intervention, comparator, outcome, study design
PO	Oral administration
PRIMA	Preliminary independent model advice
PSA	Probability sensitivity analysis
PSP	Patient support programme
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTCA	Percutaneous transluminal coronary angioplasty
Q6M	Once every six months
QALY	Quality adjusted life year
QD	Once daily
QM	Once monthly
QoL	Quality of life
QW	Once weekly
r	Exposure-adjusted incidence rate per 100 subject-years
RAL	Raloxifene
RANK	Receptor activator of nuclear factor kappa-B
RCT	Randomised controlled trial
RIS	Risedronate
RoB	Risk of bias
ROM/ROMO	Romosozumab
RR	Risk ratio/relative risk
RRR	Relative risk reduction
SAE	Serious adverse event
SC	Subcutaneous
ScHARR	School of Health and Related Research
ScHARRHUD	School of Health and Related Research Health Utilities Database
SD	Standard deviation
SE	Standard error
SERM	Selective oestrogen receptor modulator
SI	Système international (d'unités); English: International System of Units
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SMR	Standardised mortality ratio
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TH	Total hip
TLV	Tandvårds- och läkemedelsförmånsverket; English: The Swedish Dental and Pharmaceutical Benefits Agency
TRP	Teriparatide
TTO	Time trade-off
Tx	Treatment
UK	United Kingdom
USA	United States of America
VBA	Visual Basic
Vert.	Vertebral
WCO-IOF-ESCEO	World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases
WHO	World Health Organization
WTP	Willingness-to-pay
ZOL	Zoledronate

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues, Section 1.2 presents the key model outcomes, Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness (CE). Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (CE) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1 provides a summary of the key issues identified by the ERG.

Table 1.1: Summary of key issues

ID3936	Summary of issue	Report Sections
1	There is a problem with the population in the CS, with comparator populations at different risks for fracture, which means none of the comparisons are reliable	2.1 and 3.4
2	It is possible that effects of romosozumab wane after 42 months	3.2.5 and 3.6
3	The network meta-analyses (NMAs) are unreliable	3.3 and 3.4
4	It is unclear whether the company's and ERG's base-case analyses are representative for UK clinical practice	4.2.4, 5.1 and 6.2
5	Assumptions regarding persistence with osteoporosis therapies are uncertain and have a major impact on the model results	4.2.6
6	Model usability could be improved by performing calculations in the model work sheets and by significantly reducing its running time	5.3

CS = company submission; ERG = Evidence Review Group; NMA = network meta-analysis; UK = United Kingdom

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are the following:

- Persistence rates for romosozumab and alendronate,
- Excess mortality associated to fractures (ERG assumed only for hip fractures and company also after vertebral and non-hip non-vertebral (NHNV) fractures),
- Incremental fracture and daily long-term care (LTC) costs,
- Inclusion of cardiovascular (CV) adverse events (AEs) and patient support programme (PSP) costs,
- Number of General Practitioner (GP) visits per year, and
- The source of United Kingdom (UK) general population mortality rates.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival; OS) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Reducing the incidence of fractures, and
- QALYs are reduced by cardiovascular (CV) adverse events (AEs).

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments, and
- Reducing costs associated to a decreased number of fractures.

The modelling assumptions that have the greatest effect on the ICER are:

- Treatment persistence
- Treatment effect of romosozumab followed by alendronate and alendronate alone
- Utility multipliers for hip, vertebral and NHNV fracture
- Comparator choice
- Inclusion of CV AEs
- Assumed excess mortality
- Start age of the population
- Model time horizon

1.3 The decision problem: summary of the ERG's key issues

The decision problem addressed in the company submission (CS) is broadly in line with the final scope issued by NICE. However, there is a problem with the population in the CS, which means none of the comparisons are reliable (Table 1.2).

Table 1.2: Key issue 1: There is a problem with the population in the CS, with comparator populations at different risks for fracture, which means none of the comparisons are reliable

Report Section	Sections 2.1 and 3.4
Description of issue and why the ERG has identified it as important	• The population in the CS (imminent risk of a fracture, i.e. having had a MOF within the last 2 years) is narrower than the scope, which does not define "high risk" or mention a time limit, and the ARCH ITT population where some patients without any time limit were included. In the NMAs the populations in the comparator studies are diverse, but mainly include women at high risk of a fracture as in the ARCH ITT population. • The ARCH trial includes a head-to-head comparison of romosozumab vs. alendronate. Both treatments are recommended for women at high risk of a fracture. However, oral bisphosphonates (such as alendronate) are recommended for the "high risk" group and anabolic agents (such as romosozumab) are recommended for the "very high risk" group (Kanis et al. 2020). Therefore, the comparison, romosozumab vs. alendronate may not be the appropriate comparison in the very high risk subgroup.

Report Section	Sections 2.1 and 3.4
What alternative approach has the ERG suggested?	The submission should only focus on the "imminent risk" population in the ARCH trial. This population is as specified in the CS and allows a head-to-head comparison with alendronate.
What is the expected effect on the cost effectiveness estimates?	The effectiveness results used in the model are based on the NMA for the ITT population. However, the overall model is based on a different population, the imminent risk population. It would be useful if the company could add a scenario where both effectiveness data and the whole model are based on the imminent risk population from the ARCH trial.
What additional evidence or analyses might help to resolve this key issue?	Clinical expert opinion as to whether imminent risk is consistent with only high as opposed to very high risk or whether it also includes very high risk. This would provide clarity as to whether alendronate is the most appropriate comparator.
CS = company submission; ERG = Evidence Review Group; ITT = intention-to-treat; MOF = major osteoporotic fracture; NMA = network meta-analysis	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The ERG identified two major concerns with the evidence presented on the clinical effectiveness, namely that it is possible that effects of romosozumab wane after 42 months (Table 1.3) and that the network meta-analyses (NMAs) are unreliable (Table 1.4).

Table 1.3: Key issue 2: It is possible that effects of romosozumab wane after 42 months

Report Section	Sections 3.2.5 and 3.6
Description of issue and why the ERG has identified it as important	The Kaplan-Meier curves for time to first clinical fracture and time to first non-vertebral fracture show that there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first fracture up to month 42. However, the curves seem to converge again by month 48. This means that it is possible that the effects of romosozumab wane over time. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Therefore, longer term follow-up is needed to see whether the effects are maintained over time.
What alternative approach has the ERG suggested?	The economic evaluation should include a scenario where treatment waning starts at 4 years followed by a dynamic offset (linear waning) of the treatment effect. The economic evaluation should also include a scenario where the dynamic offset of the treatment effect is shorter (e.g., three years).
What is the expected effect on the cost effectiveness estimates?	In the base-case analysis, treatment effect is maintained for 5 years (60 months). After that, a dynamic offset (linear waning) of the treatment effect is assumed for another 5 years. At year 11, there is no treatment effect. An early treatment effect waning can be modelled by using larger hazard ratios. This would increase the ICER.
What additional evidence or analyses might help to resolve this key issue?	For the first scenario no additional evidence is necessary. For the second scenario the company would need to adjust the model to allow selecting different durations of the dynamic offset of the treatment effect.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 1.4: Key issue 3: The network meta-analyses (NMAs) are unreliable

Report Section	Sections 3.3 and 3.4
Description of issue and why the ERG has identified it as important	The NMAs are unreliable for the following reasons: • There was little direct evidence for comparisons for romosozumab included in any of the NMAs. • Most studies had differences in mean age, ethnicity, or rate of prevalent vertebral fractures, indicating at least a moderate RoB from effect modification. • As almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone. • Individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). • There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the risk of bias. • As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low risk of bias; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high risk of bias.
What alternative approach has the ERG suggested?	There is no alternative approach with the data available in the CS, beyond interpreting the effect estimates with due caution from the high-RoB present in almost all comparisons, with the exceptions of alendronate and placebo (which had direct evidence). To reduce bias, either of the following is possible, though would require additional data: 1. Include direct evidence from more trials of romosozumab and comparator treatments, by conducting more trials; and 2. Request individual participant data from all trials included in the NMAs and adjust for known effect modifiers. This option only decreases bias, and large biases may remain due to an inability to adjust away all effects of differences in effect modifiers between trials.
What is the expected effect on the cost effectiveness estimates?	The expected effect on the CE estimates is uncertain.

Report Section	Sections 3.3 and 3.4
What additional evidence or analyses might help to resolve this key issue?	To reduce bias, either of the following is possible, though would require additional data: 1. Include direct evidence from more trials of romosozumab and comparator treatments, by conducting more trials; and 2. Request individual participant data from all trials included in the NMAs and adjust for known effect modifiers. This option only decreases bias, and large biases may remain due to an inability to adjust away all effects of differences in effect modifiers between trials.
CE = cost effectiveness; CS = company submission; ERG = Evidence Review Group; NMA = network meta-analysis; RoB = risk of bias	

1.5 The cost effectiveness evidence: summary of the ERG's key issues

A full summary of the CE evidence review conclusions can be found in Section 6.4 of this report. The company's CE results are presented in Section 5, the ERG's summary and detailed critique are in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues in the CE evidence are discussed in Tables 1.5 to 1.7.

Table 1.5: Key issue 4: It is unclear whether of the company's and ERG's base-case analyses are representative for UK clinical practice

Report Section	Sections 4.2.4, 5.1 and 6.2
Description of issue and why the ERG has identified it as important	There is uncertainty regarding the appropriateness and relevance of the comparators included in the analyses, and how these relate to the relevant population for this assessment as described in key issue 1. For example, Kanis et al. 2020 recommended that raloxifene is given to patients at low risk of fractures, oral bisphosphonates (such as alendronate and risedronate) are given to high risk patients, and anabolic agents (such as romosozumab and teriparatide) followed by an inhibitor of bone resorption (such as oral bisphosphonates) are provided to very high risk patients.
What alternative approach has the ERG suggested?	Identify what comparators are representative of UK clinical practice in the imminent risk population. After this is done, results can be selected for the right comparators only.
What is the expected effect on the cost effectiveness estimates?	As shown with the different scenario analyses, results are likely to vary depending on the comparators selected.
What additional evidence or analyses might help to resolve this key issue?	The Committee should clarify what comparators are representative of UK clinical practice in the imminent risk population.
ERG = Evidence Review Group; UK = United Kingdom	

Table 1.6: Key issue 5: Assumptions regarding persistence with osteoporosis therapies are uncertain and have a major impact on the model results

Report Section	Section 4.2.6
Description of issue and why the ERG has identified it as important	The company's approach to model persistence is inconsistent between intervention (persistence based on trial data) and comparators (persistence based on clinical practice) and is likely to be biased in favour of the intervention. Persistence assumptions

Report Section	Section 4.2.6
	were identified as one of the most important drivers of the CE results.
What alternative approach has the ERG suggested?	The ERG estimates for persistence are consistent between intervention and comparators. The ERG also identified a more recent study (Morley et al. 2020) to estimate persistence on the comparator treatments.
What is the expected effect on the cost effectiveness estimates?	When the ERG preferred base-case assumption for persistence with alendronate is applied (without the other ERG preferred changes) to the company base-case model, the ICER increased from £16,660 to £162,391 per QALY gained.
What additional evidence or analyses might help to resolve this key issue?	The uncertainty regarding persistence with osteoporosis therapies could be resolved by a study that uses data on present-day persistence in the UK, and by further investigating to what extent it is relevant to distinguish between naïve and non-naïve patients.
CE = cost effectiveness; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; UK = United Kingdom	

Table 1.7: Key issue 6: Model usability could be improved by performing calculations in the model work sheets and by significantly reducing its running time

Report Section	Section 5.3
Description of issue and why the ERG has identified it as important	Model review would be facilitated if calculations were performed in the model worksheets, instead of being hard coded in VBA. This code was initially password protected and therefore the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. After clarification, the company provided most of the VBA code which was reviewed by the ERG. No major issues were found but, nevertheless, the ERG was not allowed to make any changes to the VBA code in the model version used to run the scenarios because this model version still contains the code used for the Fracture Risk Assessment tool (FRAX), which is confidential. Additionally, the model seems to be extremely demanding regarding the computational power needed to run within a reasonable time. This makes the validation process extra difficult. The ERG did not succeed in running any probabilistic sensitivity analysis (PSA). Some counterintuitive results were observed when teriparatide was involved as a comparator treatment. The ERG was not able to find the source for these inconsistencies, which might need further confirmation from the company.
What alternative approach has the ERG suggested?	A full evaluation of the model and the assumptions included cannot be performed without access to the VBA code within the model. The ERG would like to suggest the company conduct an analysis to estimate the minimal PSA loop sizes that would provide reliable results in a minimum running time and to re-consider the programming of the model in order to make it computationally more efficient.
What is the expected effect on the cost effectiveness estimates?	It should not impact the model results but it would facilitate model validation and usability.

What additional evidence or analyses might help to resolve this key issue?	A new model version in which the ERG is allowed to make changes in the VBA code if deemed necessary. Also, a new model version with improved running time would enable the execution of a PSA.
ERG = Evidence Review Group, FRAX = Fracture Risk Assessment tool; PSA = probabilistic sensitivity analysis; VBA = Visual Basic	

1.6 Other key issues: summary of the ERG's view

No other key issues were identified by the ERG.

1.7 Summary of the ERG's view

Table 1.8 provides the incremental results of both the company's and ERG's preferred base-cases, as well as the impact of each ERG assumption change applied individually to the company base-case. As can be seen, the ERG base-case ICER is substantially larger than the company's. The change which had the largest impact by far on the results was the use of estimates for persistence on alendronate from Morley et al. 2020, which increased the ICER to £162,391. The next largest change in results was observed when assuming a daily cost of long-term care of £67 (i.e., instead of £112), which increased the ICER by nearly £6,000 per QALY gained. All other changes had an independent impact of less than £5,000 on the ICER.

The ERG was unable to run a probabilistic sensitivity analysis (PSA) for its preferred base-case analysis. However, given the deterministic ICER and assuming that the PSA ICER would be in line with this one, the probability that romosozumab is considered cost effective at a threshold of £30,000 compared to alendronate is likely to be [redacted]%. Scenario analyses run on the ERG preferred assumptions showed that model results were most sensitive to assumed rates of persistence; however, scenarios surrounding utility multipliers, treatment effect waning, excess mortality due to fractures and inclusion of CV AEs and PSP also had large impacts on the ICER, which was very sensitive to changes in the small incremental QALYs. When various alternative comparators were included in the analysis, romosozumab was dominated by zoledronate. In this situation, the only relevant comparison was zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated.

Table 1.8: Summary of ERG's preferred assumptions and ICER

Scenario	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
Company's base-case	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	162,391
+ Excess mortality only for hip fractures	[redacted]	[redacted]	17,185
+ Daily LTC costs £67	[redacted]	[redacted]	22,476
+ Incremental fracture costs	[redacted]	[redacted]	20,398
+ CV adverse events included	[redacted]	[redacted]	19,500
+ No PSP	[redacted]	[redacted]	17,680
+ 2 GP visits per year	[redacted]	[redacted]	17,117
+ UK general population mortality 2017-2019	[redacted]	[redacted]	16,903
ERG's preferred base-case	[redacted]	[redacted]	483,750

Scenario	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
CV = cardiovascular; ERG = Evidence Review Group; GP = General Practitioner; ICER = incremental cost effectiveness ratio; LTC = long-term care; PSP = probabilistic sensitivity analysis; QALY = quality adjusted life year; UK = United Kingdom			

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Postmenopausal women with severe osteoporosis at high risk of fracture	Postmenopausal women with severe osteoporosis who are at high risk of fracture and who have: • Experienced a recent MOF within 24 months; and • Thus, are at imminent risk of another fragility fracture	• Romosozumab is not licensed for use in men, in premenopausal women or in patients without severe osteoporosis • The submission positions romosozumab for use in a population that is part of the licenced population, including women with the greatest unmet need, and for whom romosozumab is expected to provide substantial clinical benefit	The population is not in line with the NICE scope. The population described in the NICE scope is the same as the licensed population for romosozumab. However, the population in the ARCH trial is narrower in that patients should have had a previous MOF. The population in the CS is narrower again in that a patient should have had a recent (within 24 months) MOF.
Intervention	Romosozumab	Romosozumab for 12 months, followed by sequential alendronate.	Romosozumab is licensed as a 12-month course of treatment. The SmPC for romosozumab states that *"following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months"*	The intervention in the CS is romosozumab for 12 months, followed by sequential alendronate.
Comparator(s)	• Bisphosphonates (alendronic acid, risedronate sodium, ibandronic acid and zoledronic acid)	The base-case comparisons are vs. alendronate, using the head-to-head ARCH study, and vs. no active treatment. Scenario analyses are provided against all other comparators	No trials of the licensed dose of ibandronate were found to be included in the NMA for fracture outcomes, therefore comparisons could not be conducted.	The comparators are in line with the NICE scope, except for the exclusion of ibandronate.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• Non-bisphosphonates (denosumab, raloxifene and teriparatide) • No active treatment	listed in the scope, using the NMA, except ibandronic acid.		
Outcomes	• Osteoporotic fragility fracture • Bone mineral density • Mortality • Adverse effects of treatment • Health-related quality of life	In line with the final NICE scope.	In line with the final NICE scope.	The outcomes reported are in line with the NICE scope.
Economic analysis	• The reference case stipulates that the CE of treatments should be expressed in terms of incremental cost per QALY • The reference case stipulates that the time horizon for estimating clinical and CE should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared • Costs will be considered from an NHS and PSS perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent	Not reported.	Not reported.	The CE analyses were conducted according to the NICE reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	treatment technologies will be taken into account			
Subgroups to be considered	If evidence allows, subgroups based on patient characteristics that increase the risk of fracture (that is, those specified in NICE clinical guideline 146) or that affect the impact of fracture on lifetime costs and outcomes should be considered.	Not reported.	Not reported.	No subgroup analyses were performed by the company.

Based on Table 1 and pages 11 to 12 of the CS[1]

CE = cost effectiveness; CS = company submission; ERG = Evidence Review Group; MOF = major osteoporotic fracture; NICE = National Institute for Health and Care Excellence; NMA = network meta-analysis; NHS = National Health Service; PSS = Personal Social Services; QALY = quality adjusted life year; SmPC = Summary of Product Characteristics

2.1 Population

The population defined in the scope is: "*Postmenopausal women with severe osteoporosis at high risk of fracture"*.[2]

The population in the company submission (CS) is limited to *"Postmenopausal women with severe osteoporosis who are at high risk of fracture and who have: Experienced a recent major osteoporotic fracture (MOF) within 24 months; and thus, are at imminent risk of another fragility fracture"*.[1]

According to the company, the decision problem addressed in the CS is narrower than that specified in the final scope and narrower than the marketing authorisation for romosozumab (CS, Section B.1.1, page 10).[1] According to the company, the patient population in the CS *"focusses on women with the greatest unmet need, and for whom romosozumab is expected to provide substantial clinical benefit"* (CS, Section B.1.1, page 10).[1]

The population included in the ARCH trial was ambulatory postmenopausal women aged 55 to 90 years if they had at least one of the following bone mineral density (BMD) and fracture criteria:

- BMD T-score at the total hip or femoral neck of ≤ -2.50 and EITHER:
    - at least one moderate (SQ2) or severe (SQ3) vertebral fracture OR
    - at least two mild (SQ1) vertebral fractures

OR

- BMD T-score at the total hip or femoral neck of ≤ -2.00 and EITHER:
    - at least two moderate (SQ2) or severe (SQ3) vertebral fractures OR
    - a fracture of the proximal femur that occurred within three to 24 months prior to randomisation

In addition, at least one hip must have been evaluable by dual-energy x-ray absorptiometry (DXA).

Assuming that all vertebral fractures are considered major osteoporotic fractures (MOFs), the population in the CS is largely in line with the population in the main trial, the ARCH trial, in which postmenopausal women who have previously suffered a MOF have been included.[3] However, the company does explain that the ARCH population is not completely in line with the population in the CS, with the key difference being that the ARCH trial did not mandate the prior fracture to be recent, whereas the romosozumab target population (i.e. the population in the CS) defines recency of fracture as a criterion (CS, page 43).[1]

In the ARCH trial, a total of [redacted] patients had suffered a fracture within zero to 24 months before randomisation ([redacted] in the romosozumab/alendronate group; [redacted] in the alendronate alone group). Of these, [redacted] patients in the romosozumab/alendronate group and [redacted] patients in the alendronate alone group suffered a recent MOF and would be eligible for treatment with romosozumab according to the target patient population considered in the CS.

In 2019, a European marketing authorisation was granted for romosozumab. Romosozumab is indicated for the treatment of severe osteoporosis in postmenopausal women at high risk of fracture.[4] Romosozumab is contraindicated for patients with: hypersensitivity to the active substance(s) or to any of the excipients, hypocalcaemia, or a history of MI or stroke.[4]

In summary, there seem to be three relevant populations:

1. The population as described in the NICE final scope,[2] which is the same as the European marketing authorisation for romosozumab: Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is not defined;
2. The population in the ARCH trial (intention-to-treat (ITT) population):[3] Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is defined as having previously suffered a MOF; and
3. The population in the CS:[1] Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is defined as having suffered a fracture within the last two years (also referred to as 'imminent risk of fracture').

There is also a lack of clarity as to the difference between "high risk" and "very high risk". For example, Kanis et al. 2020 recommended that raloxifene is given to patients at low risk of fractures, oral bisphosphonates (such as alendronate and risedronate), are given to high risk patients, and anabolic agents (such as romosozumab and teriparatide) followed by an inhibitor of bone resorption (such as oral bisphosphonates) are provided to very high risk patients. However, it is not clear whether current clinical practice in the UK is based on these or similar recommendations. Multiple treatment guidelines are available that differ in their (wording of) recommendations and it is not clear which treatment guideline is both up-to-date and relevant for the NHS. This therefore raises the question as to whether "high" and "very high" are mutually exclusive or whether "high" includes "very high": if the former, then comparators other than alendronate might not be appropriate comparators, but if the latter then they might be.

2.2 Intervention

The intervention (romosozumab) is in line with the scope. However, romosozumab is licensed as a 12-month course of treatment. The Summary of Product Characteristics (SmPC) for romosozumab states that *"following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months"*.[4] Therefore, the intervention in the CS is *"romosozumab for 12 months, followed by sequential alendronate"* (CS, Table 1, page 11).[1]

The recommended dose of romosozumab is 210 mg, which is administered as two subcutaneous (SC) injections of 105 mg each into the abdomen, thigh or upper arm.[4] The use of romosozumab is limited to once during a lifetime (CS, page 22).[1]

According to the company, no additional tests or investigations are required prior to the administration of romosozumab (CS, page 13).[1]

2.3 Comparators

The description of the comparators in the NICE scope is as follows: *"Bisphosphonates (alendronic acid, risedronate sodium, ibandronic acid and zoledronic acid), Non-bisphosphonates (including antiresorptive agents (denosumab, raloxifene) and anabolic agents (teriparatide)), and No active treatment"*.[2]

In the CS, the base-case comparisons are vs. alendronate, using the head-to-head ARCH study, and vs. no active treatment. Scenario analyses are provided against all other comparators listed in the scope, using the network meta-analysis (NMA),, except ibandronic acid. According to the company, *"no trials*

of the licensed dose of ibandronate were found to be included in the NMA for fracture outcomes, therefore comparisons could not be conducted" (CS, Table 1, page 11).[1]

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- Osteoporotic fragility fracture
- Bone mineral density
- Mortality
- Adverse effects of treatment
- Health-related quality of life

These were all assessed in the ARCH trial. However, the ARCH trial had a median follow-up duration of 33 months, at which time 90 participants in each group had died.[3] Therefore, if romosozumab is expected to improve survival, the follow-up is insufficient to show any differences.

Regarding health-related quality of life (HRQoL), the company states that the trial data do not provide HRQoL values sensitive to decreases in HRQoL after a fracture. In addition, the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited, according to the company.[1]

2.5 Other relevant factors

According to the company, romosozumab is innovative because it *"is the only dual-acting osteoanabolic biologic, with all other treatments being antiresorptives or a single-action anabolic. Antiresorptive therapies do not directly stimulate bone formation and therefore, romosozumab provides a clear advantage over bisphosphonates by rapidly increasing bone formation on naïve bone surface resulting in rapid improvements in bone density, mass, microstructure and strength leading to superior fracture risk reductions"* [5, 6] (CS, Section B.2.11).[1]

A patient access scheme (PAS) has been proposed for romosozumab. The proposed romosozumab with PAS price is £[redacted] per monthly dose, equivalent to a percentage discount of [redacted]%. This equates to an annual cost of £[redacted] (with PAS; CS, Section B.1.2, page 13).[1]

This appraisal does not fulfil the end-of-life criteria as specified by NICE because the life expectancy of patients eligible for romosozumab is well beyond 24 months. Therefore, treatment is not indicated for patients with a short life expectancy (normally less than 24 months).

According to the company, romosozumab is only licensed for use in postmenopausal women, not men. However, *"osteoporosis is four times more likely to occur in women than men, and is prevalent in 21.8% of women (versus 6.8% of men) over the age of 50 in the UK"*[7] (CS, Section B.1.4).[1]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company conducted a systematic review (an original review and two updates) to evaluate the evidence on clinical effectiveness (efficacy and safety) of romosozumab for the treatment of osteoporosis in postmenopausal women at increased risk of fracture.[8] Section 3.1 critiques the methods of the review including: the search strategy; study inclusion criteria; data extraction; assessment of risk of bias; and data synthesis.

3.1.1 Searches

Appendix D of the CS provided details of the systematic literature searches used to identify clinical efficacy and safety evidence.[8] Database searches were conducted in August 2016, updated in March 2018, and updated again in September 2020. Summaries of the resources searched for each set of searches are provided in Tables 3.1 to 3.3.

Table 3.1: Resources searched for clinical efficacy and safety, August 2016

Resource	Host/source	Date range	Date searched
Databases			
Embase	OvidSP	1974 to 17 August 2016	18 August 2016
MEDLINE	OvidSP	1946 to August Week 2 2016	24 August 2016
MEDLINE In-Process Citations, Epub Ahead of Print & Daily Update	OvidSP	up to 24 August 2016	24 August 2016
PubMed	NLM	up to 25 August 2016	25 August 2016
CDSR	Wiley Online Library	Issue 8/August 2016	16 August 2016
CENTRAL	Wiley Online Library	Issue 7/July 2016	16 August 2016
DARE	Wiley Online Library	Issue 2/April 2015	16 August 2016
HTA Database	Wiley Online Library	Issue 3/July 2016	16 August 2016
PROSPERO	http://www.crd.york.ac.uk/PROSPERO/	Not reported	Not reported

Resource	Host/source	Date range	Date searched
GIN Library	http://www.g-i-n.net	Not reported	Not reported
Clinical Trial Registries			
ClinicalTrials.gov	https://clinicaltrials.gov	Not reported	Not reported
WHO ICTRP	http://www.who.int/ictrp/en	Not reported	Not reported
Conference proceedings			
NOF	https://www.nof.org/	2013 and 2014	26 August 2016
NOS	https://nos.org.uk/	2014	6 October 2016
WCO-IOF-ESCEO	http://www.wco-iof-esceo.org/	2013, 2014, 2015 and 2016	25 August 2016
HTA websites			
CADTH	https://www.cadth.ca/	Not reported	Not reported
EMA / CHMP	http://www.ema.europa.eu	Not reported	Not reported
NICE	http://www.nice.org.uk	Not reported	Not reported
NIHR	http://www.nets.nihr.ac.uk/	Not reported	Not reported
US Drugs @ FDA	https://www.accessdata.fda.gov/scripts/cder/daf/	Not reported	Not reported

CADTH = Canadian Agency for Drugs and Technologies in Health; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CHMP = Committee for Medicinal Products for Human Use; DARE = Database of Abstracts of Reviews of Effects; EMA = European Medicines Agency; FDA = Food & Drug Administration; GIN = Guidelines International Network; HTA = health technology assessment; ICTRP = International Clinical Trials Registry Platform; NICE = National Institute for Health and Care Excellence; NIHR = National Institute for Health Research; NLM = National Library of Medicine; NOF = National Osteoporosis Foundation; NOS = National Osteoporosis Society; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases; WHO = World Health Organization

Table 3.2: Resources searched for clinical efficacy and safety, March 2018

Resource	Host/source	Date range	Date searched
Databases			
Embase	OvidSP	1974 to 27 March 2018	28 March 2018

Resource	Host/source	Date range	Date searched
MEDLINE	OvidSP	1946 to March Week 3 2018	28 March 2018
MEDLINE In-Process Citations, Epub Ahead of Print & Daily Update	OvidSP	up to 27 March 2018	27 March 2018
PubMed	NLM	up to 28 March 2018	28 March 2018
CENTRAL	Wiley Online Library	Issue 12/ February 2018	28 March 2018
Northern Light Life Sciences Conference Abstracts	Ovid	2010 to Week 11 2018	Not reported
Clinical Trial Registries			
ClinicalTrials.gov	https://clinicaltrials.gov	Not reported	Not reported
WHO ICTRP	http://www.who.int/ictrp/en	Not reported	Not reported
Conference proceedings			
NOF	https://www.nof.org/	2013 to 2016	Not reported
NOS	https://nos.org.uk/	2014 and 2016	Not reported
WCO-IOF-ESCEO	http://www.wco-iof-esceo.org/	2013 to 2017	Not reported

CENTRAL = Cochrane Central Register of Controlled Trials; ICTRP = International Clinical Trials Registry Platform; NOF = National Osteoporosis Foundation; NOS = National Osteoporosis Society; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases; WHO = World Health Organization

Table 3.3: Resources searched for clinical efficacy and safety, September 2020

Resource	Host/source	Date range	Date searched
Databases			
Embase	Not reported	Not reported	Not reported
PubMed	Not reported	Not reported	Not reported
Cochrane	Not reported	Not reported	Not reported

ERG comment:

- The selection of databases searched was very comprehensive. Full details of the database searches including the database name, host platform and date searched, were provided.
- Conference proceedings were searched. Details of the conferences searched, URLs, and the date of the searches were provided. The search strategies or search terms used, and results were not reported in the CS.[1] In response to the request for clarification, the company explained that relevant

conference publications were identified from the Embase search and that an additional search for conference publications was conducted in Northern Light Life Sciences Conference Abstracts.[9] The search strategy used to search Northern Light Life Sciences Conference Abstracts was provided in response to the request for clarification.

- Trials registers were searched, but details of the search strategies or search terms used, dates of searches, and results were not reported in the CS. Details of the trials registers searched and the search strategies used were provided in response to the request for clarification.[9]
- Health technology assessment (HTA) organisation websites were searched, but details of the search terms used, dates of searches, and results were not reported in the CS.[1] Details of the search terms used were provided in response to the request for clarification.[9]
- Extensive use of truncation, proximity operators, synonyms, and subject headings (MeSH and EMTREE) were included in the search strategies. Cited study design search filters for randomised controlled trials (RCTs) were included. There were no language or date limits.
- Separate searches for safety data were not conducted. Ideally, a search for AEs should be carried out alongside the search for effectiveness.[10]
- Update searches were conducted in March 2018 and September 2020. Full details of the March 2018 searches were provided, but only the databases searched were provided for the September 2020 update. Details of the search strategies and results for the September 2020 update were provided in response to the request for clarification.[9] The September 2020 searches did not directly replicate the original 2016 and March 2018 searches.

3.1.2 Inclusion criteria

As stated above, the company performed a systematic review to evaluate the evidence on clinical effectiveness (efficacy and safety) of romosozumab for the treatment of osteoporosis in postmenopausal women at increased risk of fracture.[1] The original systematic review was conducted in 2016 and the two subsequent updates in 2018 and 2020.[8] The study eligibility criteria for the original and updated systematic reviews are summarised in Table 3.4 below.

Table 3.4: Eligibility criteria used in the original and updated systematic reviews of clinical effectiveness evidence

	Inclusion criteria	Exclusion criteria
Population	Studies had to include: • Postmenopausal women with osteoporosis at increased risk of fracture Where trials included a mixed population of participants where not all these inclusion criteria were fulfilled, the study was excluded unless separate data were reported for the population of interest.	Studies recruiting the following were excluded: • Women being studied for the prevention or treatment of glucocorticoid induced osteoporosis • Women with normal or unspecified BMD who have not been selected based on the presence of risk factors • Women with other indications for osteoporosis treatment e.g., Paget's disease, hypercalcaemia of malignancy, metastatic breast cancer
Interventions	The intervention of interest was romosozumab (CDP7851/AMG 785; Amgen Inc. and UCB Inc.), a monoclonal antibody that binds and	Not applicable.

	Inclusion criteria	Exclusion criteria
	inhibits sclerostin, a negative regulator of bone formation, dosed at 210 mg SC QM for 12 months for the treatment of osteoporosis.	

Comparators	Eligible comparator therapies were pharmacological therapies and those in development (in accordance with the UK, European, and US licensed indications):
	• Placebo (in accordance with NICE TAG4627)
	• Usual care e.g., vitamin D and calcium supplementation (in accordance with NICE TAG4627)
	• Antibody-based RANK ligand therapy:
	  • Denosumab (Prolia, AMG 162; Amgen Inc.)
	• Parathyroid hormone-based therapy:
	  • Teriparatide (Forteo/Forsteo; Eli Lily)
	  • Abaloparatide (BA058; Radius Health)
	• Bisphosphonates (in accordance with NICE TAG4627):
	  • Alendronate (Fosamax; Merck Sharp & Dohme; also available non-proprietary)
	  • Risedronate (Actonel; Procter & Gamble UK)
	  • Ibandronate (Boniva; Hoffman La Roche)
	  • Zoledronic acid/zoledronate (Aclasta/Reclast; Novartis)
	• Selective oestrogen receptor modulators (SERMs):
	  • Raloxifene (Evista, LY139481; Eli Lilly)
	  • Strontium ranelate (Protelos; Servier Laboratories) (*subsequently excluded)*
		The following interventions were excluded:
		• Odanacatib (Merck) – following September 2016 protocol amendment to inclusion criteria
		• Strontium ranelate (Protelos; Servier Laboratories) – following March 2018 protocol amendment to the inclusion criteria[a]
		• Combination therapies (with the exception of usual care as described above)
		• Interventions which were not administered in accordance with licensed indications
		• Interventions which were co-administered with any other therapy with the potential to augment bone unless concomitant treatments were specified in the SmPC and applied equivalently in all study arms.

Outcomes	Studies had to report the occurrence of at least one of the following fracture outcomes: • New vertebral fracture	Studies were excluded from the review if they: • Did not report at least one prespecified fracture outcome

	Inclusion criteria	Exclusion criteria
	• Clinical vertebral fracture • Non-vertebral fracture • Clinical fracture • Hip fracture Fracture outcomes were classified using the definitions provided in each specific study.	• Only reported fractures as part of the adverse event monitoring process (e.g., a BMD outcome study reporting fractures outcomes as adverse events was excluded) • Reported outcomes relating to fractures associated with major trauma (e.g., road traffic accidents). Studies that reported mixed trauma and/or non-trauma fracture, were only included if they reported separate data for relevant non-trauma fractures
Study design	To be included in the review, trials had to fulfil the following criteria:[b] • Use a parallel RCT design. This included randomised dose finding and formulation trials with either a placebo or active control arm and was not limited by study phase • Followed-up patients for at least 12 months	The following were excluded: • Systematic reviews and pooled analyses (used for reference checking purposes only and not included in the review, unless the data are not available from publications of the individual trials) • Studies based on animal models • Pre-clinical and biological studies • Narrative reviews, letters, editorials, and opinions
Language restrictions	• No restrictions for clinical effectiveness review. • English language only for review of economic evaluations, cost and resource use studies.	

Based on Table 13 of Appendix D of the CS.[8]
a Only relevant to the review update. b This was in accordance with relevant criteria from the recent HTA undertaken by NICE (ScHARR, The University of Sheffield) in March 2015 to assess TA464 - Bisphosphonates for prevention osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161).[11]
BMD = bone mineral density; CS = company submission; HTA = Health Technology Assessment; NICE = National Institute of Health and Care Excellence; QM = once monthly; RANK = receptor activator of nuclear factor kappa-B; RCT = randomised controlled trial; SC = subcutaneous; ScHARR = School of Health and Related Research; SERM = selective oestrogen receptor modulator; SmPC = Summary of Product Characteristics; TA = technology appraisal; UK = United Kingdom; US = United States (of America)

ERG comment:

Population

As outlined in Section 2.1, three relevant populations have been described. One of these is the ITT population in the ARCH trial (postmenopausal women with severe osteoporosis at high risk of fracture, the latter being defined as a previous MOF) which is used as the basis for a series of NMAs and economic modelling in the CS.[3]

We note that some placebo-controlled RCTs providing comparator arms for the NMAs recruit populations with different characteristics to those described in the ARCH trial[3] i.e., they recruit a proportion of participants without evidence of prevalent vertebral fracture at baseline. Examples include (with percentages indicating the proportion of women without prevalent vertebral fracture at baseline): two RCTs evaluating zoledronic acid (36% to 40%);[12, 13] one RCT evaluating raloxifene (75%);[14] and one RCT assessing denosumab (73%)[15]. These RCTs did not provide outcome data on subgroups defined according to presence/absence of prevalent vertebral fracture at baseline. Whilst the proportions with and without fracture at baseline were balanced across treatment groups within the individual RCTs, the populations were unlikely to be comparable to that of the ARCH trial in the context of NMA.[3]

Language restrictions

There were no language restrictions for the clinical effectiveness review and this is line with recommended good practice in SLRs.[16]

3.1.3 Critique of data extraction

In section D.2 of Appendix D of the CS, it is stated that data from each included trial were extracted into a Microsoft Excel template by a reviewer who was familiar with the subject area and validated by a second, independent reviewer.[8] The response to the clarification questions confirmed that disagreements were resolved through discussion and if necessary, by consulting a third reviewer.[9] Recommended good practice is dual, independent data extraction, particularly for outcome data.[16] In light of this, the possibility of errors within the data extraction cannot be discounted.

3.1.4 Quality assessment

Section D.2 of Appendix D explains that the risk of bias (RoB) within each included study was assessed using the Cochrane RoB tool for RCTs[8] and the company's response to the clarification questions confirmed that the original version of the tool was used.[9] Although this tool is appropriate for assessing the quality of RCTs, it is not clear why the most recent version was not used (Cochrane RoB 2).[16] One reviewer assessed the RoB and a second reviewer independently checked the assessment. Any discrepancies were resolved through consensus.[8]

3.1.5 Evidence synthesis

It was not feasible to pool the identified, eligible RCTs using direct data, pairwise meta-analysis because of differences in populations and treatment comparisons. An indirect treatment comparison was performed and this is discussed in Sections 3.3 and 3.4.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The clinical effectiveness evidence for romosozumab in severe osteoporosis in the CS is mainly based on the ARCH trial. Two other phase III clinical trials, the FRAME and STRUCTURE trials are mentioned in the CS as well. However, neither the FRAME nor STRUCTURE trial studied a patient population aligned to where the company expects romosozumab to be used in NHS clinical practice; therefore these two trials will be briefly discussed in Section 3.2.7 of this report.[17, 18] A fourth study, the BRIDGE study,[19] considered use in men, which is not part of the marketing authorisation for romosozumab; as such, no clinical effectiveness results are presented from BRIDGE in the CS. However, some data from BRIDGE are introduced in the safety section of the CS and will be discussed in Section 3.2.6 of this report.

3.2.1 Details of the included trial: the ARCH trial

The ARCH trial is a phase III, multicentre, randomised, double-blind trial, comparing romosozumab followed by alendronate vs. alendronate alone in postmenopausal women with severe osteoporosis and a fragility fracture (see Table 3.5).[3] This trial provides evidence for romosozumab in its expected position in the clinical pathway: a first-line therapy in patients who have previously suffered a MOF. Efficacy outcomes reported in the ARCH trial include incidence of clinical, vertebral, non-vertebral and hip fracture and percentage change from baseline in BMD. Data from the ARCH trial were used as the main data for the economic modelling in this submission.

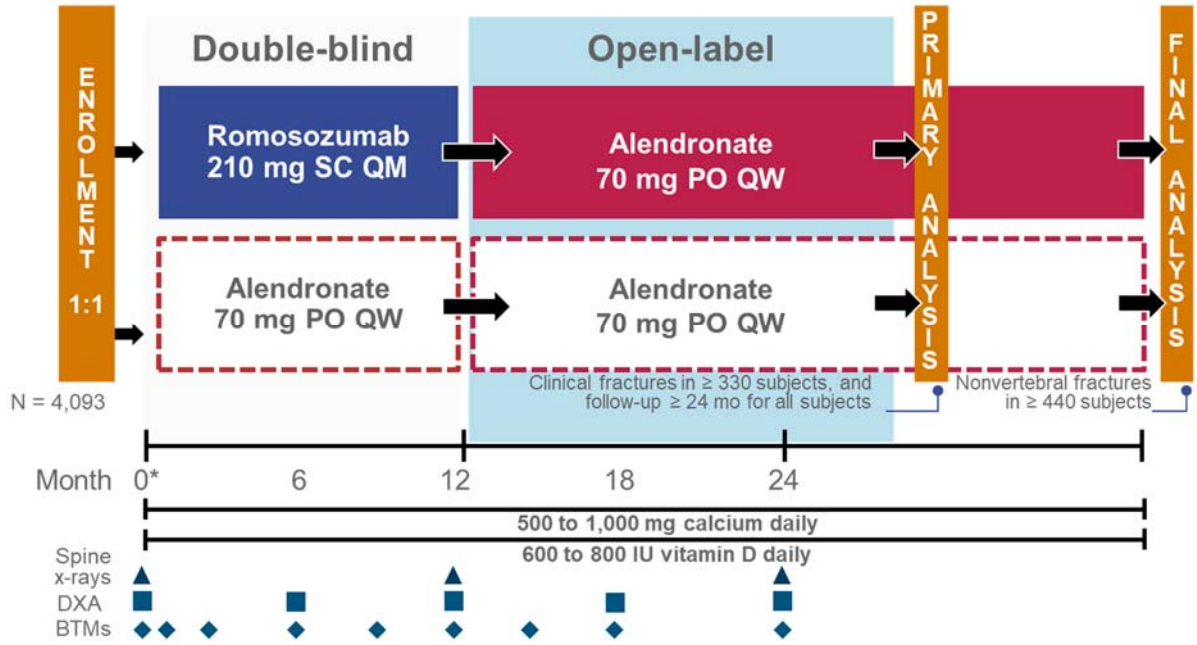
Table 3.5: Summary of methodologies for the ARCH trial

Trial number (acronym)	**NCT01631214 (ARCH)**
Study design	International, multicentre, randomised, double-blind, active-controlled, parallel-group, phase III.
Location	This study was conducted at [redacted] centres across Europe, North America, Central and South America, and Asia/Pacific, including [redacted] sites in the UK (76 UK patients out of 4,093).
Population	Ambulatory postmenopausal women aged ≥55 to ≤90 years of age at randomisation who met at least one of the following criteria: • BMD T-score of ≤–2.5 at TH or FN and either ≥1 moderate or severe vertebral fractures or ≥2 mild vertebral fractures • BMD T-score of ≤–2.0 at TH or FN and either ≥2 moderate or severe vertebral fractures, or a fracture of the proximal femur sustained three to 24 months prior to randomisation • At least one hip that could be evaluated by DXA
Duration of study	Double-blind treatment period: 12 months. Open-label period: minimum 12 months (until end of study).
Method of randomisation	Patients were randomly assigned to receive romosozumab or alendronate using IVRS. Randomisation was stratified by age (<75 years vs. ≥75 years).
Method of blinding	Double blind: patients and site staff remained blinded to the patient's original treatment assignment. Treatment assignment was only unblinded if the knowledge of the treatment was essential for the patient's further management.
Intervention(s)	Romosozumab (210 mg) QM SC for 12 months followed by open-label oral alendronate (70 mg) QW for at least 12 months (until study end).
Comparator(s)	Oral alendronate (70 mg) QW for 12 months followed by open-label alendronate (70 mg) for at least 12 months (until study end).
Permitted and disallowed concomitant medication	With the exception of the medications listed in the protocol, investigators may have prescribed any concomitant medications or treatments necessary to provide adequate supportive care.
Reported outcomes relevant to the decision problem	• Cumulative incidence of new vertebral fracture • Cumulative incidence of clinical fracture • Incidence of fractures (non-vertebral, all fractures, new or worsening vertebral, major non-vertebral, hip, MOF) • Percent change in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs

Trial number (acronym)	NCT01631214 (ARCH)
Based on CS, Tables 4 to 6, pages 29-33.[1] AE = adverse event; BMD = bone mineral density; BPI = Brief Pain Inventory; CS = company submission; DXA = dual-energy X-ray absorptiometry; EQ-5D = EuroQoL-5 Dimensions; FN = femoral neck; IVRS = interactive voice response system; LAD = limited activity days; LS = lumbar spine; MOF = major osteoporotic fracture; OPAQ-SV = Osteoporosis Assessment Questionnaire Short Version; QM = once monthly; QW = once weekly; SC = subcutaneous; TH = total hip; UK = United Kingdom	

The ARCH trial comprised the following study periods: initial screening and enrolment, double-blind treatment period, and open-label treatment period (Figure 3.1). Eligible patients were randomly assigned 1:1 to receive SC romosozumab 210 mg QM or oral alendronate 70 mg QW for the first 12 months (the double-blind period). Following this, patients received open-label oral alendronate 70 mg QW for the remainder of the study (the open-label period). Initial study drug given remained blinded until completion of the open-label period.

Figure 3.1: ARCH trial design



Based on CS, Figure 3, page 31.[1]

Footnotes: All patients received daily calcium (500 mg to 1,000 mg) and vitamin D (600 IU to 800 IU). *Patients with serum 25 (OH) vitamin D levels of ≥20 mg/mL and ≤40 ng/mL at screening received an initial loading dose of 50,000 to 60,000 IU of vitamin D. The final analysis (end-of-study) occurred when non-vertebral fracture events were confirmed for at least 440 subjects, or earlier if the primary analysis demonstrated superiority of romosozumab treatment for non-vertebral fracture risk reduction.

BTM = bone turnover markers; CS = company submission; DXA = dual-energy X-ray absorptiometry; IU = international unit; PO = oral administration; QM = once monthly; QW = once weekly; SC = subcutaneous

The ARCH trial was designed as an event-driven trial. The primary analysis for ARCH was performed after all patients had completed their month 24 visit and at least 330 patients had confirmed events of clinical fracture (composite of non-vertebral fracture and clinical vertebral fracture (a suspected vertebral fracture that is brought to medical attention and confirmed)). The median follow-up time at primary analysis was 2.7 years (33 months; interquartile range (IQR), 2.2 to 3.3). For all patients, BMD

was assessed at baseline and every 12 months at the lumbar spine, total hip and femoral neck by dual-energy X-ray absorptiometry (DXA).

The primary endpoints in the ARCH trial were the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis. Key secondary endpoints included incidence of non-vertebral fracture at primary analysis and percent change in BMD compared to baseline at months 12 and 24, at the lumbar spine, total hip, and femoral neck. Additional secondary endpoints included other fractures including hip fracture.

3.2.2 Statistical analyses of the ARCH trial

In the ARCH trial, a total of 4,093 patients were randomised to the initial treatment period, with 3,654 (89.3%) patients that completed the trial up to month 12 and 3,150 (77.0%) completed the primary analysis period. The trial population used for the analysis of outcomes in ARCH are detailed in Table 3.6.

Table 3.6: Trial populations for the ARCH trial

Analysis	NCT01631214 (ARCH)
Per protocol analysis set	Included patients in the full analysis set (for clinical and non-vertebral fracture) and the primary efficacy analysis set for vertebral fractures (for new vertebral fractures) who received active investigational products and met all of the patient eligibility criteria. Used to analyse clinical fracture, new vertebral fracture, and non-vertebral fracture through month 24, clinical and non-vertebral fracture at time of primary analysis, and non-vertebral fracture at final analysis as a sensitivity analysis.
Full analysis set	Included all randomised patients in the trial. They were analysed according to their randomised treatment assignments. This was the primary analysis set used for non-vertebral fracture, clinical fracture, clinical vertebral fracture, all fracture, major non-vertebral fracture, MOF, and hip fracture endpoints.
Primary efficacy analysis set	Included all randomised patients who had a baseline and ≥1 post-baseline evaluation of vertebral fracture at or before the timepoint of consideration. Patients were analysed according to their randomised treatment assignments. This was the primary analysis set for new, new or worsening, and multiple new or worsening vertebral fractures endpoints. Patients whose first post-baseline spinal radiograph showed no fracture on vertebra, but who had the same vertebrae at baseline were also included as it could be inferred that their baseline scores would have also reported no fracture, had they been available.
Safety analysis set	Patients who received ≥1 active dose of investigational product in the 12-month double-blind study period were included in this study set. Safety data analysis for the double-blind study period, primary analysis period, and overall study period used this safety analysis set.
Based on CS, Table 8, pages 34-35.[1] CS = company submission; MOF = major osteoporotic fracture	

A summary of the statistical tests that were used during primary analysis of ARCH, and the methods by which missing data were managed, is presented in Table 3.7. For new vertebral fractures through month 12 or month 24, and clinical and non-vertebral fractures through month 12, month 24 and to primary analysis, subgroup analyses were conducted for age, presence or absence of severe vertebral fracture at baseline, number of prevalent fractures at baseline, race, geographical region, Central/Latin

America and all regions excluding Central/Latin America, baseline lumbar spine BMD T-score, baseline total hip or femoral neck BMD T-score, Fracture Risk Assessment tool (FRAX) score and history of non-vertebral fracture at age ≥55 years. For change from baseline in BMD, subgroup analyses were conducted at month 12 and month 24 for age, geographical region, baseline BMD T-score at the lumbar spine and baseline BMD T-score at the total hip.

Table 3.7: Statistical tests for the primary analysis of ARCH

Trial number (acronym)	NCT01631214 (ARCH)
Hypothesis objective	Statistical hypothesis: 12 months treatment with romosozumab followed by alendronate is effective in reducing the incidence of a clinical fracture and new vertebral fracture in postmenopausal women with osteoporosis, compared to treatment with alendronate alone.
Statistical tests	Kaplan Meier estimates were used to summarise the cumulative incidence of fracture and a Cox proportional-hazards model stratified for age and prevalent vertebral fracture was used as a basis to assess treatment comparisons. A logistic regression model based on the primary efficacy analysis set for vertebral fractures was used to compare patient incidence of new vertebral fractures up to month 24. Adjusted odds ratio and the corresponding 95% CI were also given. To demonstrate the robustness of the primary analytical model results, additional supportive analysis was conducted including: per protocol analyses and time-to-event analysis based on full analysis set. The statistical significance for the primary and selected key secondary endpoints were controlled using sequential testing procedure to maintain the overall significance level for the study at 0.05. If both the primary endpoints were significant at the 0.05 level (2-sided), each of the following secondary DXA BMD endpoints were tested hierarchically at 0.05 (2-sided). With this procedure, formal inferential testing was performed for a step only when statistical significance was declared for all endpoints tested in previous steps. If the testing sequence stopped, the remaining endpoints in the testing sequence were not formally tested for statistical significance and the corresponding p-values were considered descriptive. The p-values for the analyses of other secondary, exploratory, and sub-study endpoints were nominal without adjusting for multiplicity. All p-values were 2-sided.
Data management, patient withdrawals	For BMD, missing data was dealt with by using LOCF. Patients who had missing data for a scheduled visit were not included in the safety data collections for that time point (no imputation). *Post hoc* analysis of vertebral fractures using a multiple-imputation method was performed for all randomly assigned patients. Observed data (excluding any imputed values) was reported through to 36 months including BMD scores at month 36.

Based on CS, Table 9, page 36.[1]
BMD = bone mineral density; CI = confidence interval; CS = company submission; DXA = dual-energy X-ray absorptiometry; LOCF = last-observation-carried-forward

ERG comment: The ERG has no particular concerns about the statistical analysis of the ARCH trial.

3.2.3 Baseline characteristics of the ARCH trial

In the ARCH trial, nearly all patients had experienced an osteoporotic fracture prior to the trial (99.1% in alendronate arm vs. 98.8% in romosozumab arm). Of the participants that were randomised to the

alendronate or romosozumab arms, a similar number had suffered non-vertebral fractures (13.4% vs. 13.2%) or vertebral fractures (25.2% vs. 27.7%), respectively, in the two years before enrolment. Participants had a mean age of approximately 74 years.[3] Baseline characteristics were comparable across both treatment groups. Key baseline demographics and clinical characteristics for the patients included in the full analysis set in ARCH are presented in Table 3.8.

Table 3.8: Baseline characteristics in the full analysis set in the ARCH trial

Characteristic	Alendronate (N=2,047)	Romosozumab (N=2,046)
Mean age, years (SD)	74.2 (7.5)	74.4 (7.5)
Age ≥75 years, no. (%)	1,071 (52.3)	1,073 (52.4)
Ethnic group, no. (%)		
Hispanic	662 (32.3)	631 (30.8)
Non-Hispanic	1,385 (67.7)	1,415 (69.2)
Geographical region, no. (%)		
Central or Eastern Europe or Middle East	798 (39.0)	835 (40.8)
Latin America	727 (35.5)	674 (32.9)
Western Europe, Australia, or New Zealand	264 (12.9)	269 (13.1)
Asia-Pacific or South Africa	216 (10.6)	213 (10.4)
North America	42 (2.1)	55 (2.7)
Mean BMI (SD)	25.36 (4.42)	25.46 (4.41)
Mean BMD T-score (SD)		
Lumbar spine	–2.99 (1.24)	–2.94 (1.25)
Total hip	–2.81 (0.67)	–2.78 (0.68)
Femoral neck	–2.90 (0.50)	–2.89 (0.49)
Previous osteoporotic fracture at ≥45 years of age, no. (%)	2,029 (99.1)	2,022 (98.8)
Prevalent vertebral fracture, no. (%)	1,964 (95.9)	1,969 (96.2)
Grade of most severe vertebral fracture[a]		
Mild	73 (3.6)	68 (3.3)
Moderate	570 (27.8)	532 (26.0)
Severe	1,321 (64.5)	1,369 (66.9)
Previous non-vertebral fracture at ≥45 years of age, no. (%)	770 (37.6)	767 (37.5)
Previous hip fracture, no. (%)[b]	179 (8.7)	175 (8.6)
Mean FRAX MOF risk (SD)	20.0 (10.1)	20.2 (10.2)
Median serum β-CTX, ng/l (IQR)[c]	230.0 (137.0–388.0)	276.0 (166.0–407.0)
Medium serum P1NP, µg/l (IQR)[c]	44.7 (32.7–64.4)	50.6 (37.5–64.7)
Median 25-hydroxyvitamin D, ng/ml (IQR)	27.6 (24.0–34.2)	28.4 (24.0–34.8)

Based on CS, Table 7, pages 33-34.[1]

[a] The grade of the most severe fracture was assessed with the use of the Genant grading scale.[6] [b] Previous hip fracture excludes pathologic or high-trauma hip fracture. [c] Data shown are for the 266 patients (128 in the

Characteristic	Alendronate (N=2,047)	Romosozumab (N=2,046)
alendronate group and 138 in the romosozumab group) who enrolled in the biomarker sub-study and who had measurements of bone-turnover markers both at baseline and at one or more visits after baseline. β-CTX = Beta-C-Terminal Telopeptide of Type 1 Collagen; BMD = bone mineral density; BMI = body mass index; CS = company submission; FRAX = Fracture Risk Assessment tool; IQR = interquartile range; MOF = major osteoporotic fracture; P1NP = Procollagen Type 1 N-Telopeptide; SD = standard deviation		

3.2.4 Risk of bias assessment of the ARCH trial

The RoB of the ARCH trial will be discussed in Section 3.3.4 of this report, together with the STRUCTURE and FRAME trials.

3.2.5 Efficacy results of the ARCH trial

The results from the ARCH trial presented in the CS describe those that were detailed in the ARCH clinical study report (CSR) and were determined using the standard methodology of last observation carried forward (LOCF) imputation for missing data, as pre-specified in the statistical analysis plan. The data more recently presented in the peer-reviewed New England Journal of Medicine publication regarding fractures and BMD were determined using a multiple imputation for the missing data,[3] as requested by the journal. As this does not reflect the original pre-specified analyses for the ARCH trial, the company did not include these results in their submission. The ERG asked the company to clarify whether there were any differences in estimates of effect between the two methods of imputation, and to describe how any differences between these analyses could affect the CE estimate (Clarification Letter, Question A13).[9] According to the company, the methodology used to derive the clinical effectiveness for vertebral fractures in the ARCH trial had no bearing on the results:

- Hazard ratio (HR) for new vertebral fractures at 12 months were 0.63, 95% CI 0.47 to 0.85 and 0.64, 95% CI 0.46 to 0.89) using multiple imputation and LOCF, respectively; and
- HR for new vertebral fractures at 24 months were 0.52 (0.40-0.66) and 0.50 (0.38-0.66) using multiple imputation and LOCF, respectively

Therefore, the results below will be based on the data presented in the CS.

In the ARCH trial, romosozumab/alendronate statistically significantly reduced the incidence of new vertebral fractures at month 24, meeting its primary endpoint. Patients in the romosozumab/alendronate arm had a 50% lower relative risk of vertebral fractures compared to patients on alendronate alone over 24 months (Table 3.9).[20] Additionally, a statistically significantly lower proportion of patients experienced a clinical fracture (non-vertebral fracture and clinical vertebral fracture) at the time of primary analysis in the romosozumab/alendronate group compared to alendronate alone, meeting the other primary endpoint.[20] Patients treated with romosozumab had a statistically significantly greater increase in BMD from baseline compared to alendronate (adjusted P<0.001), which was maintained until month 36 (Table 3.9).[20]

Table 3.9: Summary of clinical effectiveness results from ARCH

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
Primary outcomes			
Incidence of new vertebral	147/1834 (8.0%)	74/1825 (4.1%)	RR= 0.50 [redacted] (0.38, 0.66)

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
fracture at 24 months			
Incidence of clinical fracture at time of primary analysis (median 33 months)	266/2047 (13.0%)	198/2046 (9.7%)	HR= 0.73 [redacted] (0.61, 0.88)
Key secondary end points			
Incidence of non-vertebral fracture at the time of the primary analysis	217/2047 (10.6)	178/2046 (8.7)	HR= 0.81 [redacted] (0.66, 0.99)
BMD Outcomes: N, LS Mean (SE) – Mean Difference (95% CI)			
BMD at the lumbar spine at 12 months	1718, 5.0 ([redacted])	1722, 13.7 ([redacted])	MD = 8.7 (8.31, 9.09)
BMD at the lumbar spine at 24 months	1577, 7.2 ([redacted])	1571, 15.3 ([redacted])	MD = 8.1 (7.58, 8.57)
BMD at the lumbar spine at 36 months	1597, 7.8 ([redacted])	1593, 15.2 ([redacted])	MD = 7.4 (6.84, 7.89)
BMD at the total hip at 12 months	1781, 2.8 ([redacted])	1781, 6.2 ([redacted])	MD = 3.3 (3.03, 3.60)
BMD at the total hip at 24 months	1627, 3.5 ([redacted])	1622, 7.2 ([redacted])	MD = 3.8 (3.42, 4.10)
BMD at the total hip at 36 months	1653, 3.5 ([redacted])	1653, 7.2 ([redacted])	MD = 3.7 (3.29, 4.02)
BMD at the femoral neck at 12 months	1781, 1.7 ([redacted])	1781, 4.9 ([redacted])	MD = 3.2 (2.90, 3.54)
BMD at the femoral neck at 24 months	1627, 2.3 ([redacted])	1622, 6.0 ([redacted])	MD = 3.8 (3.40, 4.14)
BMD at the femoral neck at 36 months	1653, 2.4 ([redacted])	1653, 6.0 ([redacted])	MD = 3.6 (3.18, 3.97)

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
Other secondary end points			
Incidence of new vertebral fracture at 12 months	85/1703 (5.0%)	55/1696 (3.2%)	RR = 0.64 [redacted]; (0.46, 0.89)
Incidence of clinical fracture at 12 months	110/2047 (5.4)	79/2046 (3.9)	HR = 0.72 [redacted]; (0.54, 0.96)
Incidence of clinical fracture at 24 months	[redacted]	[redacted]	[redacted]
Incidence of non-vertebral fractures at 12 months	95/2047 (4.6)	70/2046 (3.4)	HR = 0.74 [redacted]; (0.54, 1.01)
Incidence of non-vertebral fractures at 24 months	[redacted]	[redacted]	[redacted]
Incidence of clinical vertebral fracture at 12 months	18/2047 (0.9)	10/2046 (0.5)	HR = 0.56 [redacted]; (0.26, 1.22)
Incidence of clinical vertebral fracture at 24 months	44/2047 (2.1)	18/2046 (0.9)	HR = 0.41 [redacted]; (0.24, 0.71)
Incidence of hip fractures at 12 months	22/2047 (1.1)	14/2046 (0.7)	HR = 0.64 [redacted]; (0.33, 1.26)
Incidence of hip fractures at 24 months	[redacted]	[redacted]	[redacted]
Incidence of hip fractures at primary analysis	66/2047 (3.2)	41/2046 (2.0)	HR = 0.62 [redacted]; (0.42, 0.92)
Incidence of major nonvertebral fractures at 12 months	88/2047 (4.3)	59/2046 (2.9)	HR = 0.67 [redacted]; (0.48, 0.94)
Incidence of major	196/2047 (9.6)	146/2046 (7.1)	HR = 0.73 [redacted]; (0.59, 0.90)

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
nonvertebral fractures at primary analysis			
Incidence of major osteoporotic fractures at 12 months	85/2047 (4.2)	61/2046 (3.0)	HR = 0.72 [redacted]; (0.52, 1.01)
Incidence of major osteoporotic fractures at primary analysis	209/2047 (10.2)	146/2046 (7.1)	HR = 0.68 [redacted]; (0.55, 0.84)
Incidence of all osteoporotic fractures at 12 months	189/2047 (9.2)	134/2046 (6.5)	HR = 0.71 [redacted]; (0.57, 0.88)
Incidence of all osteoporotic fractures at primary analysis	392/2047 (19.1)	266/2046 (13.0)	HR = 0.65 [redacted]; (0.56, 0.76)

Based on CS, Section B.2.6, pages 38-43; CSR, Section 10.[1, 20]

[a] Values < 1 for RR favour romosozumab; based on the Mantel-Haenszel method adjusted for age strata, baseline total hip BMD T-score (≤-2.5, >-2.5), and presence of severe vertebral fracture at baseline; [b] SE represents the standard error of log (risk ratio); c Hazard ratio < 1 favours romosozumab; The HR estimate is based on Cox proportional hazards model adjusting for age strata, baseline total hip BMD T-score, and presence of severe vertebral fracture at baseline.

BMD = bone mineral density; CI = confidence interval; CS = company submission; HR = hazard ratio; MD = mean difference; RR = risk ratio; SE = standard error

As shown in Figure 3.2 there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first clinical fracture by month 12. At the time of primary analysis, patients treated with romosozumab/alendronate had a lower cumulative incidence of clinical fracture (9.7%) compared to the alendronate/alendronate group (13.0%) (nominal and adjusted P<0.001). This equated to a 27% lower relative risk of clinical fracture in the romosozumab/alendronate group than alendronate alone, meeting the co-primary endpoint for the ARCH trial.

ERG comment: Although the curves diverge from months zero to 42, they seem to converge again by month 48. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Longer term follow-up is needed to see whether the effects are maintained over time.

Figure 3.2: Kaplan-Meier curves for time to first clinical fracture

Based on CS, Figure 7, page 40.[1]
Footnote: Risks presented are based on a LOCF method for patients with missing fracture status. For Kaplan-Meier curves in the time-to-event analysis, data from patients who withdrew or reached the end of the reporting period without having a fracture were carried forward from the last observation time.
CS = company submission; LOCF = last observation carried forward; N = number of patients randomised; n = number of patients at risk for event at time point of interest

Similarly, patients treated with romosozumab showed a visible separation in time to non-vertebral fracture at month 12 compared to alendronate-treated patients, which was maintained for the duration of the study (Figure 3.3).[3]

ERG comment: Similar as in Figure 3.2, the curves in Figure 3.3 diverge from months 0 to 42 and seem to converge again by month 48. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Longer term follow-up is needed to see whether the effects are maintained over time.

Figure 3.3: Kaplan-Meier curves for time to first non-vertebral fracture

Based on CS, Figure 9, page 41.[1]

CS = company submission; N = number of subjects randomised; n = number of subjects at risk for event at time point of interest

ERG comment: Overall, results of the ARCH trial are favourable for romosozumab. Both primary outcomes (the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis) are met and most fracture results significantly favour romosozumab over alendronate. In addition, all BMD outcomes significantly favour romosozumab over alendronate. However, the graphs for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3), seem to indicate that the effectiveness of romosozumab over alendronate becomes less after 42 months; longer term follow-up is needed to see whether the effects are maintained over time.

3.2.5.1 Health-related quality of life

[redacted] in health-related quality of life (HRQoL) were observed between treatment groups in the ARCH trial.[1] According to the company, "this was to be expected because the HRQoL data were collected at predetermined, discrete time points irrespective of fracture occurrence during the trial and always related to one of the investigated treatments. Therefore, the trial data do not provide HRQoL values sensitive to decrease in HRQoL after a fracture, and are hence expected to underestimate the potential HRQoL gain with treatment".[1] The company also points out that it is *"important to note that the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited"*.[1] [redacted] By preventing fragility fractures, romosozumab (and alendronate) are expected to prevent future HRQoL decrements resulting from a fracture, according to the company.

3.2.6 Adverse events

3.2.6.1 Adverse events in the ARCH trial

The incidences of AEs and serious adverse events (SAEs) were similar overall in the ARCH trial between the two treatment groups during the 12-month double-blind period, and cumulative incidences were similar between the two groups during the primary analysis period (Table 3.10). In the first 12 months, injection-site reactions (mostly mild in severity) were reported in more patients receiving romosozumab (90 of 2,040 patients (4.4%)) than in those receiving alendronate (53 of 2,014 patients (2.6%)).

However, more people in the romosozumab group experienced adjudicated serious CV AEs during the double-blind period, with 50 patients (2.5%) in the romosozumab group and 38 (1.9%) in the alendronate group reporting these events (odds ratio, 1.31; 95% confidence interval (CI) 0.85 to 2.00). A total of 16 patients (0.8%) in the romosozumab group and 6 (0.3%) in the alendronate group reported cardiac ischemic events (odds ratio, 2.65; 95% CI, 1.03 to 6.77), and 16 patients (0.8%) in the romosozumab group and seven (0.3%) in the alendronate group reported cerebrovascular events (odds ratio, 2.27; 95% CI, 0.93 to 5.22) (Table 3.10).

Table 3.10: Adverse events in the ARCH trial

Event	Month 12: Double-Blind Period		Primary Analysis: Double-Blind and Open-Label Period*	
	Alendronate (N = 2,014)	Romosozumab (N = 2,040)	Alendronate to Alendronate (N = 2,014)	Romosozumab to Alendronate (N = 2,040)
	number of patients (percent)			
Adverse event during treatment	1,584 (78.6)	1,544 (75.7)	1,784 (88.6)	1,766 (86.6)
Back pain†	228 (11.3)	186 (9.1)	393 (19.5)	329 (16.1)
Nasopharyngitis†	218 (10.8)	213 (10.4)	373 (18.5)	363 (17.8)
Serious adverse event	278 (13.8)	262 (12.8)	605 (30.0)	586 (28.7)
Adjudicated serious cardiovascular event‡	38 (1.9)	50 (2.5)	122 (6.1)	133 (6.5)
Cardiac ischemic event	6 (0.3)	16 (0.8)	20 (1.0)	30 (1.5)
Cerebrovascular event	7 (0.3)	16 (0.8)	27 (1.3)	45 (2.2)
Heart failure	8 (0.4)	4 (0.2)	23 (1.1)	12 (0.6)
Death	12 (0.6)	17 (0.8)	55 (2.7)	58 (2.8)
Noncoronary revascularisation	5 (0.2)	3 (0.1)	10 (0.5)	6 (0.3)
Peripheral vascular ischemic event not requiring revascularisation	2 (<0.1)	0	5 (0.2)	2 (<0.1)
Death	21 (1.0)§	30 (1.5)	90 (4.5)§	90 (4.4)
Event leading to discontinuation of trial regimen	64 (3.2)	70 (3.4)	146 (7.2)	133 (6.5)
Event leading to discontinuation of trial participation	27 (1.3)	30 (1.5)	43 (2.1)	47 (2.3)
Event of interest¶				
Osteoarthritis‖	146 (7.2)	138 (6.8)	268 (13.3)	247 (12.1)
Hypersensitivity	118 (5.9)	122 (6.0)	185 (9.2)	205 (10.0)
Injection-site reaction**	53 (2.6)	90 (4.4)	53 (2.6)	90 (4.4)
Cancer	28 (1.4)	31 (1.5)	85 (4.2)	84 (4.1)
Hyperostosis††	12 (0.6)	2 (<0.1)	27 (1.3)	23 (1.1)
Hypocalcaemia	1 (<0.1)	1 (<0.1)	1 (<0.1)	4 (0.2)
Atypical femoral fracture‡	0	0	4 (0.2)	2 (<0.1)
Osteonecrosis of the jaw‡	0	0	1 (<0.1)	1 (<0.1)

Based on Saag et al. 2017.[3]

* Incidence rates at the time of the primary analysis were cumulative and included all events in the double-blind and open-label period (to February 27, 2017) in patients who received at least one dose of open-label

Event	Month 12: Double-Blind Period		Primary Analysis: Double-Blind and Open-Label Period*	
	Alendronate (N = 2,014)	Romosozumab (N = 2,040)	Alendronate to Alendronate (N = 2,014)	Romosozumab to Alendronate (N = 2,040)
	number of patients (percent)			

alendronate; † Shown are events that occurred in 10% or more of the patients in either group during the double-blind period; ‡ Serious cardiovascular adverse events were adjudicated by the Duke Clinical Research Institute, and potential cases of osteonecrosis of the jaw and atypical femoral fracture were adjudicated by independent committees. Cardiovascular deaths include fatal events that were adjudicated as being cardiovascular-related or undetermined (and, therefore, possibly cardiovascular-related); § One patient had a non–treatment-related serious adverse event of pneumonia that was incorrectly flagged as death in the primary analysis snapshot and was not included in the analysis of fatal events; ¶ Events of interest were those that were identified by prespecified Medical Dictionary for Regulatory Activities search strategies; ‖ Prespecified events that were reported under osteoarthritis were osteoarthritis, spinal osteoarthritis, exostosis, arthritis, polyarthritis, arthropathy, monoarthritis, and interspinous osteoarthritis; ** The most frequent adverse events of injection-site reactions (occurring in >0.1% of the patients) in the romosozumab group during the double-blind period included injection-site pain (in 1.6% of the patients), erythema (1.3%), pruritus (0.8%), haemorrhage (0.5%), rash (0.4%), and swelling (0.3%); †† Prespecified events reported under hyperostosis were exostosis (mostly reported as heel spurs), lumbar spinal stenosis, spinal column stenosis, cervical spinal stenosis, enostosis, extra skeletal ossification, and vertebral foraminal stenosis.

3.2.6.2 Pooled adverse events from seven romosozumab studies

The safety and tolerability of romosozumab was evaluated in a programme including seven clinical trials, exposing more than 7,500 patients to romosozumab. The safety data presented in this section is a pooled analysis of the studies listed in Table 3.11, which includes the BRIDGE trial in men.

Table 3.11: Overview of studies included in the pooled safety analysis

Study	Design	Number of patients included in safety set
FRAME	Multicentre, international, randomised, double-blind, placebo-controlled, parallel-group, Phase III	Safety analysis set (n=7,157)
ARCH	Multicentre, international, randomised, double-blind, active-controlled, Phase III	Safety analysis set (n=4,054)
NCT00896532	Dose-ranging, randomised, placebo- and active controlled in women with low BMD	Safety analysis set (n=410)
NCT01992159	Dose-ranging, placebo-controlled in Japanese postmenopausal women with osteoporosis	Safety analysis set (n=252)
STRUCTURE	Multicentre, international, randomised, open-label, active-controlled, parallel-group, Phase III	Safety analysis set (n=432)
NCT02016716	Placebo-controlled, noninferiority study of romosozumab 70 vs. 90 mg/mL in postmenopausal women with osteoporosis	Safety analysis set (n=294)
BRIDGE	Multicentre, international, randomised, double-blind, placebo-controlled, Phase III	Safety analysis set (n=244) Included the male osteoporosis population

Across the pooled safety analysis set, which included the studies outlined in Table 3.11, the incidence of treatment-emergent adverse events (TEAEs) was similar in patients treated with romosozumab compared to the control group (Tables 3.12 and 3.13); the control included patients treated with placebo, alendronate and teriparatide across the clinical trial programme; exposure-adjusted incidence rate per 100 patient years: [redacted] events per 100 years (romosozumab) vs. [redacted] events per 100 years (control). Treatment related SAEs leading to discontinuation of study drug were also comparable (Table 3.12; exposure-adjusted incidence rate per 100 patient years of [redacted] in both the control group and romosozumab 210 mg QM group).

In the pooled studies, [redacted]% of patients treated with 210 mg QM romosozumab reported a serious TEAE, compared to [redacted]% of patients in the control group (Table 3.12). The most common serious TEAE reported was pneumonia ([redacted]% romosozumab 210 mg QM-treated patients vs. [redacted]% control-treated patients).

Table 3.12: Summary of exposure-adjusted incidence rate of treatment emergent adverse events (osteoporosis safety analysis set)

	All Studies (Including ARCH)		
	Control[a] (N=[redacted]) n (r)	**Romosozumab 210 mg QM[b] (N=[redacted]) n (r)**	**Romosozumab Total[c] (N=[redacted]) n (r)**
All treatment-emergent adverse events			
All TEAEs	[redacted]	[redacted]	[redacted]
Serious AEs	[redacted]	[redacted]	[redacted]
Leading to discontinuation of investigational product	[redacted]	[redacted]	[redacted]
Fatal AEs*	[redacted]	[redacted]	[redacted]
Treatment-related treatment-emergent adverse events[d]			
Treatment-related TEAEs	[redacted]	[redacted]	[redacted]
Serious AEs	[redacted]	[redacted]	[redacted]
Leading to discontinuation of investigational product	[redacted]	[redacted]	[redacted]
Fatal AEs	[redacted]	[redacted]	[redacted]

Based on CS, Table 14, page 55.[1]

* Alendronate-treated subject 14248015041 had a fatal non-treatment-related serious AE of pneumonia that had an incorrect death flag in the primary analysis snapshot and was not included in the exposure-adjusted incidence rate of fatal events; a Includes placebo from Studies FRAME (12 months), NCT00896532 (24 months), NCT01992159 (12 months), BRIDGE (12 months), and NCT02016716 (6 months), alendronate from Studies NCT00896532 (12 months) and ARCH (12 months), and teriparatide from studies NCT00896532 (12 months), and STRUCTURE (12 months); b Includes Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months), NCT01992159 (12 months), 20110142 (12 months), BRIDGE (12 months), and NCT02016716 (6 months); c Includes romosozumab QM and Q3M from Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months, all data), NCT01992159 (12 months), ARCH (12 months), BRIDGE (12 months), and NCT02016716 (6 months); d Includes only events for which the investigator indicated there was a reasonable possibility they may have been caused by investigational product.
AE = adverse event; CS = company submission; QM = every month; QW = every week; r = exposure-adjusted incidence rate per 100 subject-years; TEAE = treatment-emergent adverse event

Table 3.13: Exposure-adjusted incidence rate of most frequent (≥5.0 per 100 subject-years in total romosozumab or integrated control groups) adverse events by preferred term (osteoporosis safety analysis set)

Preferred term*	All Studies (Including ARCH)		
	Control[a] (N=[redacted]) n (r)	Romosozumab 210 mg QM[b] (N=[redacted]) n (r)	Romosozumab Total[c] (N=[redacted]) n (r)
Number of patients reporting treatment-emergent AEs	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]
Pain in extremity	[redacted]	[redacted]	[redacted]
Fall	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]
Osteoarthritis	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]
Viral upper respiratory tract infection	[redacted]	[redacted]	[redacted]

Source: CS, Table 15, page 56.[1]

* Preferred terms are sorted by descending order of the exposure-adjusted incidence rate in the total romosozumab group and control group and coded using Medical Dictionary for Regulatory Activities version 19.1; a Includes placebo from Studies FRAME (12 months), NCT00896532 (24 months), NCT01992159 (12 months), BRIDGE (12 months), and NCT02016716 (6 months), alendronate from Studies NCT00896532 (12 months) and ARCH (12 months) and teriparatide from Studies NCT00896532 (12 months), and STRUCTURE (12 months); b Includes Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months), NCT01992159 (12 months), ARCH (12 months), BRIDGE (12 months), and NCT02016716 (6 months); c Includes romosozumab QM and Q3M from Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months, all data), NCT01992159 (12 months), 20110142 (12 months), BRIDGE (12 months), and NCT02016716 (6 months).

AE = adverse event; CS = company submission; QM = every month; QW = every week; r = exposure-adjusted incidence rate per 100 subject-years; TEAE = treatment-emergent adverse event

3.2.7 Included studies: Supporting evidence

According to the company, the clinical effectiveness evidence for romosozumab in severe osteoporosis is provided from three phase III clinical trials: ARCH, FRAME and STRUCTURE. A fourth study, BRIDGE, considered use in men, which is not part of the marketing authorisation for romosozumab; as such, no clinical effectiveness results are presented from BRIDGE in the CS.[1] However, some data from BRIDGE are introduced in the pooled safety analysis (see Section 3.2.6 of this report).

The ARCH trial has been discussed in the sections above. Neither the FRAME nor STRUCTURE trials studied a patient population aligned to where romosozumab is expected to be used in NHS clinical practice. In addition, STRUCTURE was also not designed to evaluate fracture outcomes.[17, 18] Therefore, the FRAME and STRUCTURE trials will only be minimally discussed in this section of the ERG report.

Table 3.14: Supporting evidence

Study	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
Study design	International, multicentre, randomised, double-blind, placebo-controlled, parallel-group, Phase III.	International, multicentre, randomised, open-label, active-controlled, parallel-group, phase III.
Population	• Postmenopausal women with osteoporosis • Aged 55–90 years	• Postmenopausal women with osteoporosis transitioning from bisphosphonate therapy • Aged 55–90 years • Prior fragility fracture
Intervention(s)	Romosozumab (210 mg) QM SC for 12 months followed by open-label denosumab (60 mg) SC Q6M for 24 months (until study end).	Romosozumab (210 mg) QM SC for 12 months.
Comparator(s)	Placebo QM SC for 12 months followed by open-label denosumab (60 mg) Q6M SC for 24 months (until study end).	Daily SC teriparatide (20 µg) for 12 months.
Reported outcomes relevant to the decision problem	• Incidence of a new vertebral fracture • Cumulative incidence of non-vertebral fracture, major non-vertebral fracture, clinical fracture, hip fracture, new or worsening vertebral fracture, MOF and multiple new or worsening vertebral fractures • Percent change from baseline in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs	• Percent change from baseline in BMD at LS, TH, and FN • Finite element analysis of the hip[a] • AEs

Based on CS, Table 4, page 29.[1]
AE = adverse event; BMD = bone mineral density; BPI = Brief Pain Inventory; CS = company submission; EQ-5D-5L = EuroQoL-5 Dimensions-5 Levels Health Survey; FN = femoral neck; LAD = limited activity days; LS = lumbar spine; MOF = major osteoporotic fracture; OPAQ-SV = Osteoporosis Assessment Questionnaire Short Version; SC = subcutaneous; TH = total hip; Q6M = once every six months; QM = once monthly; TH = total hip

3.2.7.1 The FRAME Study

The FRAME study demonstrated statistically significant reductions in new vertebral fractures for romosozumab compared with placebo at 12 months follow-up (relative risk reduction (RRR): 73%; absolute risk reduction (ARR): 1.30%; adjusted P<0.001). Similarly, patients in the romosozumab/denosumab arm showed a statistically significant 75% reduction in RR of new vertebral fracture compared to the placebo/denosumab arm (ARR: 1.89%; incidence of new vertebral fracture: 0.6% vs. 2.5%; 95% CI: 60 to 84; adjusted P<0.001) at 24 months follow-up.[17] Romosozumab also reduced the risk of clinical fracture (non-vertebral and clinical vertebral fracture) by 36% compared with placebo at 12 months follow-up (adjusted and nominal P=0.008) and to 33% at 24 months follow-up (adjusted P=0.096, nominal P=0.002).[17]

3.2.7.2 The STRUCTURE Study

The STRUCTURE study provides BMD and estimated bone strength data comparing romosozumab and teriparatide in a population with severe osteoporosis and who received an oral bisphosphonate before transitioning to the bone-forming agent. In the STRUCTURE study, the mean percentage change

from baseline up to month 12 in BMD at the total hip was 3.2% higher (95% CI: 2.7 to 3.8; adjusted P<0.0001) in the romosozumab group (2.6%, 95% CI: 2.2 to 3.0) compared to teriparatide (–0.6%, 95% CI –1.0 to –0.2).[18]

3.2.8 Ongoing studies

Three ongoing Post-Authorization Safety Studies (PASS) in the European Union (EU), one in the United States of America (USA) and another in South Korea are proposed to evaluate adherence to the risk minimisation measures in the romosozumab SmPC; to evaluate potential differences in serious cardiovascular AEs between romosozumab and currently-available therapies in real-world conditions; and to evaluate potential difference in serious infections between romosozumab and currently-available therapies in real-world conditions, respectively. The studies will use a multi-database approach with routinely collected data and are expected to last for a period of six years. The company is also aiming to conduct a study to assess the efficacy and safety of romosozumab in Chinese patients.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company conducted NMAs to compare the efficacy of romosozumab and romosozumab/alendronate and other bisphosphonates (alendronate, risedronate, ibandronate, zoledronate), teriparatide, denosumab and raloxifene. The ARCH, FRAME and STRUCTURE studies contributed information for the direct comparisons between romosozumab and romosozumab/alendronate with alendronate, teriparatide and placebo. Other studies comparing comparator treatments with placebo and other comparator treatments were found using the systematic review described in Section 3.1.

Five distinct outcomes were considered in the NMAs: 1) new vertebral fractures at 12, 24 and 36 months, 2) non-vertebral fractures at 12, 24 and 36 months, 3) hip fractures at 12, 24 and 36 months, lumbar spine BMD at each study's latest timepoint, 4) total hip BMD at each study's latest timepoint, and 5) femoral neck BMD at each study's latest timepoint. For fracture outcomes, results were available both for the ITT population (base-case) and the EU label population; in this report, we will focus on the ITT population results only.

The inclusion and exclusion criteria of the NMAs are shown in Table 3.15.

Table 3.15: The inclusion and exclusion criteria of the NMAs

PICOS criteria	Inclusion criteria	Exclusion criteria
Population	Postmenopausal women with osteoporosis	Did not report on the population of interest.
Interventions or Comparators	Studies comparing at least two interventions of interest (plus background therapy, defined as calcium supplements and/or vitamin D): • Placebo • Romosozumab (210 mg SC QM) • Romosozumab & Alendronate (ROMO & ALN) - 210 mg SC QM & 70 mg QW • Raloxifene (60 mg oral QD) • Alendronate (10 mg oral QD or 70mg oral QW) • Risedronate (5 mg oral QD or 35mg oral QW) • Zoledronate (5 mg IV yearly) • Denosumab (60 mg SC twice yearly)	Did not compare at least two relevant interventions.

	• Teriparatide (20 µg SC QD)* • Abaloparatide (80 mg SC QD) • Ibandronate* (150 mg oral QM)	

Outcomes	Studies reporting appropriate data for one of the following outcomes.
	• Fracture outcomes at 12, 24 and 36 months:
	  • New vertebral fracture
	  • Nonvertebral fracture
	  • Hip fracture
	• BMD outcomes (percentage change at the latest time point available from each trial):
	  • Femoral neck
	  • Lumbar spine
	  • Total hip
		Did not report any relevant outcomes or did not report appropriate data (e.g., RR but no 95% CrI, SD or SE).

Based on CS, Table 22 of Appendix D.[8] * Ibandronate was included only in the BMD outcomes. ** One trial (i.e., Hadji et al. 2012) reported on a teriparatide dose of 20 µg SC QW. ALN = alendronate; BMD = bone mineral density; CrI = credible interval; CS = company submission; IV = intravenous; PICOS = population, intervention, comparator, outcome, study design, QD = once daily; QM = once monthly; QW = once weekly; ROMO = romosozumab; RR = relative risk; SC = subcutaneous; SD = standard deviation; SE = standard error		

3.3.1 Details of the trials included in the NMAs

Different studies were included in each network for each outcome and timepoint depending on the data available, though there were similarities across networks. Networks for all fracture outcomes at 12 months used ARCH[21] for the direct comparison between romosozumab and alendronate and FRAME[22] for the direct comparison between romosozumab and placebo. Networks for fractures at 24 and 36 months used ARCH for the direct comparison between romosozumab/alendronate and alendronate. Therefore, for fracture outcomes, only indirect evidence is available for comparisons of romosozumab and romosozumab/alendronate with comparator treatments other than alendronate and placebo (at 12 months). Most studies in the NMAs for fracture outcomes compared a comparator treatment with placebo, meaning consistency cannot be assessed for most comparisons. This is because inconsistency is assessed by comparing direct and indirect comparisons of treatments, which requires a loop in a network (the simplest being a triangle, with direct evidence linking three treatments). As the vast majority of comparisons between romosozumab and comparator treatments in all NMAs only have indirect evidence, inconsistency cannot be assessed.

Networks for all BMD outcomes used ARCH[21] for the direct comparison between romosozumab and alendronate, FRAME[22] for the direct comparison between romosozumab and placebo and STRUCTURE[18] for the direct comparison between romosozumab and teriparatide. Therefore, for BMD outcomes, only indirect evidence is available for comparisons of romosozumab and romosozumab/alendronate with comparator treatments other than alendronate, teriparatide and placebo. There were more comparisons with comparator treatments other than placebo in the BMD NMAs, meaning both direct and indirect evidence if available, and so consistency could be checked for more comparisons.

Tables 3.16 and 3.17 show a list of the comparator treatments and timepoints available for each outcome for studies included in at least one network for fracture and BMD outcomes, respectively.

Table 3.16: Studies included in the NMAs of fracture outcomes

Trial/Study	Intervention	Comparator 1	Comparator 2	Included in ITT analysis	Included in EU label-matched analysis	New vertebral timepoints	Non-vertebral timepoints	Hip timepoints
ACTIVE trial[23]	Abaloparatide	Placebo	Teriparatide	Yes	Yes	24	24	24
ARCH trial[21]	Romosozumab	Alendronate	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Bai et al. 2013[12]	Zoledronate	Placebo	NA	Yes	Yes	24		24
Chao et al. 2013[24]	Zoledronate	Placebo	NA	Yes	Yes	NA	12, 36	12, 36
Dursun et al. 2001[25]	Alendronate	Placebo	NA	Yes	Yes	12	NA	NA
EVA trial	Alendronate	Placebo	NA	Yes	Yes	NA	NA	12
FIT I + II trial[26]	Alendronate	Placebo	NA	Yes	Yes	NA	36	
FIT I trial[27]	Alendronate	Placebo	NA	Yes	Yes	24, 36	NA	24, 36
FOSIT trial[28]	Alendronate	Placebo	NA	Yes	Yes	NA	12, 24	NA
FRAME trial[22]	Romosozumab	Placebo	NA	Yes	Yes	12	12	12
FREEDOM trial[15]	Denosumab	Placebo	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Hadji et al. 2012[29]	Teriparatide	Risedronate	NA	Yes	Yes	24	24	24
HORIZON-PFT trial[13]	Zoledronate	Placebo	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Liberman et al. 1995[30]	Alendronate	Placebo	NA	Yes	Yes	36	NA	NA
Liu et al. 2004[31]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
Lufkin et al. 1998[32]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
MORE trial[33]	Raloxifene	Placebo	NA	Yes	Yes	12, 24, 36	NA	NA
Morii et al. 2003[14]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
Neer et al. 2001[34]	Teriparatide	Placebo	NA	Yes	Yes	24	12, 24	24
ROSE trial[35]	Alendronate	Zoledronate	NA	Yes	Yes	NA	NA	NA
RUTH trial[36]	Raloxifene	Placebo	NA	Yes	Yes	NA	12, 24, 36	12, 24, 36
Silverman et al. 2008 (93)[37]	Raloxifene	Placebo	NA	Yes	Yes	36	36	NA

Trial/Study	Intervention	Comparator 1	Comparator 2	Included in ITT analysis	Included in EU label-matched analysis	New vertebral timepoints	Non-vertebral timepoints	Hip timepoints
VERO trial[38]	Teriparatide	Risedronate	NA	Yes	Yes	12, 24	12, 24	24
VERT MN trial (EU analysis)[39]	Risedronate	Placebo	NA	Yes	Yes		12, 24, 36	36
VERT-MN trial (AUS+EU analysis)[39]	Risedronate	Placebo	NA	Yes	Yes	12, 24, 36	NA	NA
VERT-MN trial (NAm analysis)[40]	Risedronate	Placebo	NA	Yes	Yes	12, 36	36	36
ZONE trial[41]	Zoledronate	Placebo	NA	Yes	Yes	24	NA	NA

Based on Table 24 of Appendix D of the CS.[8]
*Patients switched to alendronate after 24 months. **Patients switched to denosumab after 12 months.
Dosing schedules: Placebo, romosozumab (210 mg SC QM), raloxifene (60 mg oral QD), alendronate (10 mg oral QD or 70 mg oral QW), risedronate (5 mg oral QD or 35 mg oral QW), zoledronate (5 mg IV yearly), denosumab (60 mg SC twice yearly), teriparatide (20 µg SC QD (QW for Hadji et al. 2012)), abaloparatide (80 mg SC QD).
AUS = Australia; CS = company submission; EU = European Union; ITT = intention-to-treatment; IV = intravenous; NA = not applicable; NAm = North America; NMA = network meta-analysis; QD = once daily; QM = once monthly; QW = once weekly; SC = subcutaneous

Table 3.17: Studies included in the NMAs of BMD outcomes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
ACTIVE trial[23]	Abaloparatide	Teriparatide	Placebo		Yes	Yes	Yes
Adami et al. 1995[42]	Alendronate	Placebo			Yes	Yes	Yes
Adami et al. 2008[43]	Raloxifene	Placebo			Yes	No	Yes
Aki et al. 2004[44]	Alendronate	Placebo			Yes	No	Yes
Amgen 20010223[45]	Denosumab	Alendronate	Placebo		Yes	Yes	No
ARCH[21]	Romosozumab	Alendronate			Yes	Yes	Yes
DATA[46]	Denosumab	Teriparatide			Yes	Yes	Yes
DECIDE[47]	Denosumab	Alendronate			Yes	Yes	Yes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
DEFEND[48]	Denosumab	Placebo			Yes	Yes	Yes
Dursun et al. 2001[25]	Alendronate	Placebo			Yes	No	Yes
EFFECT[49]	Raloxifene	Alendronate			Yes	No	No
EFFECT international[50]	Alendronate	Raloxifene			Yes	Yes	Yes
EUROFORS[51]	Teriparatide	Raloxifene	Placebo		Yes	Yes	Yes
FACT[52]	Alendronate	Risedronate			Yes	Yes	Yes
FACTS1[53]	Alendronate	Risedronate			Yes	Yes	Yes
Fogelman et al. 2000[54]	Risedronate	Placebo			Yes	No	Yes
FOSIT[28]	Alendronate	Placebo			Yes	Yes	Yes
FRAME[22]	Romosozumab	Placebo			Yes	Yes	Yes
Grey et al. 2010[55]	Zoledronate	Placebo			Yes	Yes	No
Hadji et al. 2012[29]	Risedronate	Teriparatide			Yes	Yes	Yes
HORIZON[13]	Zoledronate	Placebo			Yes	Yes	Yes
Johnell et al. 2002[56]	Raloxifene	Alendronate	Placebo		Yes	No	Yes
Liberman et al. 1995[30]	Alendronate	Placebo			Yes	No	Yes
McClung et al. 2009[57]	Ibandronate	Placebo			No	Yes	Yes
McClung et al. 2014[58]	Romosozumab	Teriparatide	Alendronate	Placebo	Yes	Yes	Yes
Miller et al. 2016[59]	Denosumab	Zoledronate			Yes	Yes	No
MOTION[60]	Ibandronate	Alendronate			Yes	Yes	No
NCT00132808[61]	Zoledronate	Placebo			Yes	Yes	Yes
NCT00353080[61]	Risedronate	Placebo			Yes	No	No
NCT00398606[62]	Alendronate	Placebo			Yes	No	Yes
Neer et al. 2001[34]	Teriparatide	Placebo			Yes	Yes	Yes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
OCEAN[63]	Alendronate	Placebo			Yes	No	No
Recknor et al. 2013[64]	Denosumab	Ibandronate			Yes	Yes	Yes
Reid et al. 2011[65]	Zoledronate	Placebo			Yes	No	No
Roux et al. 2013[66]	Denosumab	Risedronate			Yes	Yes	Yes
Silverman et al. 2008[37]	Raloxifene	Placebo			Yes	Yes	No
SPIMOS[67]	Ibandronate	Placebo			Yes	Yes	No
STAND[68]	Denosumab	Alendronate			Yes	Yes	No
STRUCTURE[18]	Romosozumab	Teriparatide			Yes	Yes	Yes
Tan et al. 2016[69]	Zoledronate	Alendronate			Yes	Yes	Yes
Tucci et al. 1996[70]	Alendronate	Placebo			Yes	Yes	Yes
Um et al. 2017[71]	Raloxifene	Alendronate	Placebo		Yes	No	Yes

Source: Table 25 of Appendix D of the CS.[8]
BMD = bone mineral density; CS = company submission

3.3.2 Statistical analyses of the NMAs

For the fracture outcomes, RRs were used to estimate the relative effectiveness of all treatments, based on the number of participants in each treatment group in each study and the number of participants developing fractures by each timepoint. For BMD outcomes, mean differences (MDs) with 95% credible intervals (CrIs) were used to estimate the relative effectiveness of treatments. Some studies were missing data for the specified timepoints (12, 24 and 36 months), and were included if there were other informative timepoints, e.g. for new vertebral fractures, the ACTIVE study[23] had results at 18 months comparing abaloparatide and teriparatide, which was included in the 24-month NMA. Additionally, data from FRAME was only used at 12 months, as after 12 months, all patients in FRAME switched to denosumab.

The NMAs were conducted in a Bayesian framework: binary Bayesian models were used for fracture outcomes and shared parameter Bayesian models were used for BMD outcomes. Non-informative priors were used for all analyses. Both fixed and random effects models were presented for fracture outcomes, but only random effects models were presented for BMD outcomes due to high levels of heterogeneity observed in previous NMAs. All NMAs were run with 50,000 iterations after a burn-in of 30,000 iterations. An additional 50,000 iterations were run if the data were not sufficient converged after the initial 50,000 iterations, based on NMA diagnostic. All presented results converged.

Homogeneity was assessed using the I^2 statistic, using threshold values to indicate little (zero to 40%), moderate (30% to 60%), substantial (50% to 90%) and considerable (75% to 100%) heterogeneity. Consistency was assessed using the Bucher method, taking a P value of <0.05 as significant inconsistency, though no further action was taken in the presence of inconsistency. Baseline characteristics were compared to assess similarity of included studies, including mean age, the proportion of subjects with prevalent fracture, and mean BMD. Publication bias was not assessed.

Results were presented as tables comparing all comparator treatments, as ranks for all comparator treatments (the percentage chance of having the top, second, third rank etc.), and as forest plots showing the effectiveness of comparator treatments relative to romosozumab or romosozumab/alendronate.

3.3.3 Baseline characteristics of the trials in the NMAs

Table 3.18 details the intervention and comparator treatments for all trials included in any of the fracture NMAs, along with the outcomes and timepoints for which there were data.

Table 3.18: Trial details for all trials in any NMA of fracture outcomes

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
ACTIVE trial[23]	Abaloparatide	Placebo, Teriparatide		■			■			■	
ARCH trial[21]	Romosozumab	Alendronate	■	■	■	■	■	■	■	■	■
Bai et al. 2013[12]	Zoledronate	Placebo		■						■	
Chao et al. 2013[24]	Zoledronate	Placebo				■		■	■		■

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
Dursun et al. 2001[25]	Alendronate	Placebo	■								
EVA trial	Alendronate	Placebo							■		
FIT I + II trial[26]	Alendronate	Placebo						■			
FIT I trial[27]	Alendronate	Placebo		■	■					■	■
FOSIT trial[28]	Alendronate	Placebo				■	■				
FRAME trial[22]	Romosozumab	Placebo	■			■			■		
FREEDOM trial[15]	Denosumab	Placebo	■	■	■	■	■	■	■	■	■
Hadji et al. 2012[29]	Teriparatide	Risedronate		■			■			■	
HORIZON -PFT trial[13]	Zoledronate	Placebo	■	■	■	■	■	■	■	■	■
Liberman et al. 1995[30]	Alendronate	Placebo			■						
Liu et al. 2004[31]	Raloxifene	Placebo	■								
Lufkin et al. 1998[32]	Raloxifene	Placebo	■								
MORE trial[33]	Raloxifene	Placebo	■	■	■						
Morii et al. 2003[14]	Raloxifene	Placebo	■								
Neer et al. 2001[34]	Teriparatide	Placebo		■		■	■			■	
ROSE trial[35]	Alendronate	Zoledronate									
RUTH trial[36]	Raloxifene	Placebo				■	■	■	■	■	■
Silverman et al. 2008 (93)[37]	Raloxifene	Placebo			■			■			
VERO trial[38]	Teriparatide	Risedronate	■	■		■	■			■	
VERT MN trial (EU analysis)[39]	Risedronate	Placebo				■	■	■			■
VERT-MN trial	Risedronate	Placebo	■	■	■						

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
(AUS+EU analysis)[39]			■	■	■						
VERT-MN trial (NAm analysis)[40]	Risedronate	Placebo	■		■			■			■
ZONE trial[41]	Zoledronate	Placebo		■							

Based on Table 24 of Appendix D of the CS.[8]
CS = company submission

The company did not provide information for patient characteristics for included trials providing non-romosozumab evidence in any of the NMAs, though this information is crucial for determining whether there is a RoB in any individual comparison within an NMA. For NMAs to be unbiased, effect modifiers must be balanced across all included studies. This is particularly true if the treatment comparisons only include indirect evidence, as checking for inconsistency between direct and indirect evidence (for example, from unbalanced effect modifiers) is impossible.

The ERG has compiled a table showing the patient characteristics for all trials included in any NMA of fracture outcomes, Table 3.19. All data is taken from the original study reports reference by the company, and includes the inclusion/exclusion criteria, mean age, ethnicity and prevalence of vertebral fractures at baseline.

Table 3.19: Patient characteristics for all trials included in any NMA of fracture outcomes

Trial/study	Patient characteristics
ACTIVE trial[23]	Inclusion criteria: Postmenopausal women aged 49 to 86 years were eligible if they had BMD by dual energy x-ray absorptiometry T score of less than or equal to −2.5 and greater than −5.0 at the lumbar spine or femoral neck together with radiologic evidence of at least two mild vertebral fractures or at least one moderate vertebral fracture or history of a low-trauma fracture of the forearm, humerus, sacrum, pelvis, hip, femur, or tibia within the past 5 years. Women older than 65 years who met fracture criteria but had a T score of less than or equal to −2.0 and greater than −5.0 were eligible. Women older than 65 years were eligible without fracture criteria if either BMD T score was less than or equal to −3.0 and greater than −5.0. Eligibility required normal serum values for calcium, intact parathyroid hormone, phosphorus, and alkaline phosphatase and a 25-hydroxyvitamin D level of greater than 15 ng/mL (37.5 nmol/l (SI conversion, multiply by 2.496)). Exclusion criteria: Women were excluded if they had more than four mild, moderate, or any severe vertebral fractures (consistent with definitions described by Genant et al), fewer than two evaluable lumbar vertebrae, or if hip BMD was unevaluable. Participants were ineligible if they had evidence of metabolic bone disease or malabsorption or were taking any medications that would interfere with bone metabolism. Women were also excluded if they used bisphosphonates for more than 3 months in the past 5 years or denosumab within the past year. Women with a history of osteosarcoma were also excluded. Mean age: 69 years Ethnicity: White (80%); Asian (16%); Black or African American (3%); Other (1%) Prevalent vertebral fracture: 24%
ARCH trial[21]	Inclusion criteria: Ambulatory postmenopausal women 55 to 90 years of age who met at least one of the following criteria were eligible: a BMD T score of –2.5 or less at the total hip or femoral neck and either one or more moderate or severe vertebral fractures or two or more mild vertebral fractures; or a BMD T score of –2.0 or less at the total hip or femoral neck and either two or more moderate or severe vertebral fractures or a fracture of the proximal femur sustained 3 to 24 months before randomisation. Exclusion criteria: Patients with severe osteoporosis, an inability to take alendronate oral tablets or contraindications to alendronate, including a glomerular filtration rate below 35 ml per minute per 1.73 m^2 of body-surface area. Mean age: 74 years Ethnicity: Hispanic (32%); non-Hispanic (68%) Prevalent vertebral fracture: 96%
Bai et al. 2013[12]	Inclusion criteria: For inclusion in the study women with a primary diagnosis of osteoporosis had to be postmenopausal, have a BMD T-score <=2.5 at the femoral neck but no evidence of vertebral fractures, or a BMD T-score ≤1.5 with radiological diagnosis of two or more vertebral fractures. Exclusion criteria: (i) patients with secondary osteoporosis or other diseases known to affect bone metabolism; (ii) patients taking sodium fluoride, parathyroid hormone, anabolic steroids or growth hormone within six months of study entry, or systemic corticosteroids within

Trial/study	Patient characteristics
	12 months of study entry; (iii) patients with malignant, hepatic and renal diseases; and (iv) a serum calcium concentration of >11.0 mg/dl and untreated hypocalcaemia. Mean age: 57 years Ethnicity: Not stated (study conducted in one hospital in China) Prevalent vertebral fracture: 61%
Chao et al. 2013[24]	Inclusion criteria: Female patients diagnosed with osteoporosis. Exclusion criteria: Patients with secondary osteoporosis or other diseases which were known to affect bone metabolism were excluded. Patients taking anabolic steroids, sodium fluoride, and parathyroid or growth hormone within 6 months were also excluded. Patients who had malignant neoplasm, serum calcium more than 11.0 mg/dl, or untreated hypocalcaemia were also excluded. Mean age: 55 years Ethnicity: Not stated (study conducted in two hospitals in China) Prevalent vertebral fracture: 55%
Dursun et al. 2001[25]	Inclusion criteria: Postmenopausal women with a BMD of two SDs or more below the young adult mean at either the posteroanterior lumbar spine or the femoral neck. Exclusion criteria: Women with a documented history of drug or alcohol abuse, or with evidence from physical examination, laboratory tests or radiography of any bone metabolism disorder. Exclusion criteria also included active GI or liver disease, renal failure, renal calculi, treatment with specific therapy for osteoporosis, treatment with systemic corticosteroid therapy, malignancy, disorder of calcium metabolism and lumbar vertebrae abnormalities preventing the evaluation of BMD. Mean age: 61 years Ethnicity: Not stated (study conducted in one hospital in Turkey) Prevalent vertebral fracture: Not stated
FIT I + II trial[26]	Inclusion criteria: Women aged 55 to 80 years who had been post-menopausal for at least 2 years and had femoral neck BMD of 0.68 g/cm^2 or less. Exclusion criteria: Women with recent peptic ulcers or ulcers that required hospitalisation, dyspepsia requiring daily treatment, significant renal or hepatic dysfunction, medical problems that precluded three years of participation, severe malabsorption, blood pressure exceeding 210 mmHg systolic or 105 mmHg diastolic, MI within 6 months, unstable angina, hypothyroidism, hyperthyroidism, or hyperparathyroidism. Women taking oestrogen or calcitonin within the preceding six months or bisphosphonates or sodium fluoride (>1 mg/d) at any time were also excluded. Mean age: 68 years Ethnicity: Not stated (study conducted in 11 hospitals in the USA)

Trial/study	Patient characteristics
	Prevalent vertebral fracture: 0%
FIT I trial[27]	Inclusion criteria: Women aged 55 to 81 years who had been post-menopausal for at least 2 years and had femoral neck BMD of 0.68 g/cm^2 or less. Exclusion criteria: Women with recent peptic ulcers or ulcers that required hospitalisation, dyspepsia requiring daily treatment, significant renal or hepatic dysfunction, medical problems that precluded 3 years of participation, severe malabsorption, blood pressure exceeding 210 mmHg systolic or 105 mmHg diastolic, MI within 6 months, unstable angina, hypothyroidism, hyperthyroidism, or hyperparathyroidism. Women taking oestrogen or calcitonin within the preceding 6 months or bisphosphonates or sodium fluoride (>1 mg daily for 2 weeks or longer) at any time were also excluded. Mean age: 71 years Ethnicity: Not stated (study conducted in 11 hospitals in the USA) Prevalent vertebral fracture: 100%
FOSIT trial[28]	Inclusion criteria: Women eligible for study participation had been postmenopausal for at least 3 years, were not older than 85 years, and had BMD of the lumbar spine (L2–4) at least two standard deviations (SD) below the mean for mature, premenopausal women. Eligible patients were otherwise in good health and were between 20% below and 50% above ideal body weight as defined in the Metropolitan Life Insurance Company Height and Weight Table. Exclusion criteria: Excluded from participation were women with metabolic bone disease other than postmenopausal osteoporosis; disturbed parathyroid or thyroid function; major GI disease (for example, peptic ulcer or malabsorption) within the year before enrolment or use of a drug to inhibit gastric acid secretion for >2 weeks within 3 months of study entry; MI within the year prior to enrolment; uncontrolled hypertension or untreated angina; significantly impaired renal function (serum creatinine >150 mmol/l); or evidence of significant end organ disease. Also excluded were women who had received a bisphosphonate or fluoride (>8 mg/day) during the previous 6 months; oestrogen (except vaginal 43 times per week), ipriflavone or calcitonin during the previous 4 months; or any anabolic steroid, glucocorticoid or progestin for >2 weeks within the previous 6 months. Participants could not be receiving any medications that might alter bone or mineral metabolism, including vitamin A in excess of 10,000 U/day, vitamin D in excess of 1,000 U/day, anticonvulsants or phosphate-binding antacids. Finally, at least three vertebrae from L1 to L4 had to be evaluable by DXA to determine BMD in this region. Mean age: 63 years Ethnicity: Not stated (patients were from 153 centres in 34 countries in Europe, Latin America, Australia, Canada, South Africa and China) Prevalent vertebral fracture: Not stated
FRAME trial[22]	Inclusion criteria: Ambulatory postmenopausal women, 55 to 90 years of age, with a T score of −2.5 to −3.5 at the total hip or femoral. Exclusion criteria: Women who had a history of hip fracture, any severe or more than two moderate vertebral fractures, a history of metabolic bone disease or conditions affecting bone metabolism, osteonecrosis of the jaw, a 25-hydroxyvitamin D level of less than 20 ng

Trial/study	Patient characteristics
	per millilitre, current hypercalcemia or hypocalcaemia, or recent use of drugs affecting bone metabolism (within defined washout periods). Mean age: 71 years Ethnicity: Hispanic (40%); non-Hispanic (60%) Prevalent vertebral fracture: 18%
FREEDOM trial[15]	Inclusion criteria: Women between the ages of 60 and 90 years with a BMD T score of less than −2.5 at the lumbar spine or total hip were eligible for inclusion. Exclusion criteria: Women were excluded if they had conditions that influence bone metabolism or had taken oral bisphosphonates for more than 3 years. If they had taken bisphosphonates for less than 3 years, they were eligible after 12 months without treatment. Women were also excluded if they had used intravenous bisphosphonates, fluoride, or strontium for osteoporosis within the past 5 years; or parathyroid hormone or its derivatives, corticosteroids, systemic hormone-replacement therapy, selective oestrogen-receptor modulators, or tibolone, calcitonin, or calcitriol within 6 weeks before study enrolment. Mean age: 72 years Ethnicity: Not stated (the trial was conducted in Western Europe, Eastern Europe, Latin America, North America, Australia, and New Zealand) Prevalent vertebral fracture: 24%
Hadji et al. 2012[29]	Inclusion criteria: Women ≥45 years of age and at least 2 years postmenopausal were eligible if they had a history of back pain for ≥2 months before screening that was likely, in the opinion of the investigator, to be caused by osteoporotic vertebral fracture, despite conservative analgesic treatment; a baseline mean pain score of at least 4.0 on the numeric rating scale during the week before randomisation; lumbar spine, femoral neck, or total hip BMD T-score of ≤−2; and a minimum of one moderate vertebral fracture. Exclusion criteria: Exclusion criteria included diseases affecting bone metabolism other than osteoporosis; elevated serum calcium values, abnormal serum thyroid-stimulating hormone, parathyroid hormone, or 25-hydroxyvitamin D levels; imminent need for kyphoplasty or vertebroplasty; and evidence of significant pathology related to back pain which would make the interpretation of the back pain related to an osteoporotic vertebral fracture difficult, based on investigator assessment. Mean age: 71 years Ethnicity: Caucasian (80%); East Asian (0.4%); Hispanic (18%); Native American (0.4%), African Descent (0.8%) Prevalent vertebral fracture: 90%
HORIZON-PFT trial[13]	Inclusion criteria: Postmenopausal women between the ages of 65 and 89 years were eligible for inclusion if they had a BMD T score of −2.5 or less at the femoral neck, with or without evidence of existing vertebral fracture, or a T score of −1.5 or less, with radiologic evidence of at least two mild vertebral fractures or one moderate vertebral fracture.

Trial/study	Patient characteristics
	Exclusion criteria: Ineligibility criteria included any previous use of parathyroid hormone or sodium fluoride, use of anabolic steroids or growth hormone within 6 months before trial entry or oral or intravenous systemic corticosteroids within 12 months, and any previous use of strontium. Patients with a serum calcium level of more than 2.75 mmol per litre or less than 2.00 mmol per litre were ineligible, as were patients with a calculated creatinine clearance of less than 30.0 ml per minute at either of two baseline visits or urine dipstick results of more than 2+ for protein, without evidence of contamination or bacteriuria. Mean age: 73 years Ethnicity: Not stated (the trial was conducted in Western Europe, Eastern Europe, Latin America, North America, Oceania, and Asia) Prevalent vertebral fracture: 63%
Liberman et al. 1995[30]	Inclusion criteria: Women who were 45 to 80 years old and postmenopausal (≥5 years since menopause) with osteoporosis (defined as a BMD of the lumbar spine that was at least 2.5 SD below the mean value in premenopausal white women) were eligible for participation. Exclusion criteria: We excluded women with other causes of osteoporosis (e.g., treatment with glucocorticoids) or other disorders of bone and mineral metabolism (e.g., vitamin D deficiency, Paget's disease, or hyperparathyroidism); active peptic ulcer disease, abnormal renal function (serum creatinine level, > 1.5 mg per decilitre (130 mmol per litre)), or abnormal hepatic function; abnormalities of the lumbar spine precluding the assessment of BMD at a minimum of three lumbar vertebrae or a history of hip fracture; or any prior treatment with bisphosphonates or treatment within the preceding 12 months with oestrogen, progestin, calcitonin, fluoride, or an anabolic steroid. Mean age: 64 years Ethnicity: Not stated (the trial was conducted in the United States, Australia, Canada, Europe, Israel, Mexico, New Zealand, and South America) Prevalent vertebral fracture: 21%
Liu et al. 2004[31]	Inclusion criteria: Postmenopausal women between 50 and 80 years, who were free of severe or chronically disabling conditions, had their last menstrual period at least 2 years before the beginning of the study, and had a T-score for femoral neck or lumbar spine BMD measurements ≤2.5. Exclusion criteria: Known, suspected or history of carcinoma of the breast or oestrogen-dependent neoplasia, history of cancer within the previous 5 years, history of deep vein thrombosis, requirement of high-dose heparinization (>7500 U/d), bone disorders except for osteoporosis, treatment with any drug affecting bone metabolism, acute or chronic liver disease (bilirubin >34 umol/l, alanine transaminase >100 U/l, or alkaline phosphatase >300 U/l), impaired kidney function (serum creatinine >177 umol/l), or abnormal uterine bleeding of an unknown origin. Mean age: 65 years Ethnicity: Not stated (study conducted in three hospitals in China) Prevalent vertebral fracture: <=18%

Trial/study	Patient characteristics
Lufkin et al. 1998[32]	Inclusion criteria: Women with postmenopausal osteoporosis. Subjects were eligible if they were in good health except for osteoporosis, free of any serious acute or chronic medical condition that might affect bone or calcium metabolism, fully ambulatory, between the ages of 45 and 75 years, and postmenopausal (no menses for 5 years or levels of serum oestradiol <73 pmol/l and serum follicle-stimulating hormone (FSH) >30 IU/l). Exclusion criteria: Specific exclusion criteria included patients with a history of deep venous thrombosis, thromboembolic disorders, or cerebral vascular accident, also patients with a history of cancer within the previous 5 years, except for superficial skin cancer. Mean age: 68 years Ethnicity: Not stated (study conducted in two hospitals in the USA) Prevalent vertebral fracture: Not stated
MORE trial[33]	Inclusion criteria: Women who were at least 2 years postmenopausal, and who had osteoporosis, defined by BMD T-score of -2.5 or less and/or the presence of radiographically apparent vertebral fracture. Exclusion criteria: Not stated. Mean age: 74 years Ethnicity: Not stated Prevalent vertebral fracture: 37%
Morii et al. 2003[14]	Inclusion criteria: Women who were two or more years postmenopausal and no older than 80 years. All participants were Japanese who had osteoporosis, defined as L2-L4 BMD T-score of at least 2.5 SDs below the young adult mean and had a diagnosis consistent with the criteria for the diagnosis of osteoporosis in Japan. Exclusion criteria: Women were excluded from participation in the study if they had experienced bone disease other than primary osteoporosis, severe postmenopausal symptoms requiring oestrogen replacement therapy, history of or suspected breast carcinoma, any history of other cancer within the previous 5 years, except for excised superficial lesions; abnormal uterine bleeding, a history of deep venous thrombosis or thromboembolic disorders, as determined by evaluation of the participant questionnaire; endocrinologic disorders requiring pharmacologic therapy, acute or chronic hepatic disorder, with impaired kidney function (serum creatinine >225 lmol/l or >2.5 mg/dl); recent history of kidney stones; untreated malabsorption syndromes; or consumed an excess of alcohol or abused drugs. Participants were also excluded if, in the opinion of the investigator, they had pathologic fractures or if satisfactory evaluation of DXA could not be obtained due to X-ray findings. Patients were excluded if they had taken androgen, calcitonin, or bisphosphonate within the previous 6 months; been taking systemic oestrogen and progestin for up to one cycle (28 days) within the previous 6 months, or any systemic use within the previous 2 months; been taking the active form of vitamin D3, vitamin K2, or ipriflavone within the previous 3 months; been receiving fluoride therapy for more than 3 months during the previous 2 years; undergone systemic corticosteroid therapy for more than 1 month within the past year; or taken antiseizure drugs or pharmacologic doses of vitamin D. Participants who participated

Trial/study	Patient characteristics
	in other clinical trials within 4 months before registration or who had participated in any other clinical trial of raloxifene hydrochloride were also excluded. Mean age: 65 years Ethnicity: Japanese (100%) Prevalent vertebral fracture: 26%
Neer et al. 2001[34]	Inclusion criteria: Women were eligible for enrolment if they were ambulatory, if a period of at least five years had elapsed since menopause, and if they had at least one moderate or two mild atraumatic vertebral fractures on radiographs of the thoracic and lumbar spine, and an ambulatory status. For women with fewer than two moderate fractures, an additional criterion for enrolment was a value for BMD of the hip or lumbar spine that was at least one SD below the mean value in normal premenopausal white women (age range, 20 to 35 years). Exclusion criteria: Women with illnesses that affect bone or calcium metabolism, urolithiasis within the preceding 5 years, impaired hepatic function, a serum creatinine concentration exceeding 2 mg per decilitre (177 µmol per litre), or alcohol or drug abuse, as well as women who had taken drugs that alter bone metabolism within the previous 2 to 24 months (depending on the drug) were excluded. Mean age: 70 years Ethnicity: White (99%) Prevalent vertebral fracture: 100%
ROSE trial[35]	Inclusion criteria: Women aged 55 to 90 years who were considered postmenopausal based on either spontaneous amenorrhea or following surgical bilateral oophorectomy or after hysterectomy with serum FSH >20 IU/l and oestradiol <10 pg/ml. Eligible patients also had an increased risk of fracture, based on DXA T-score ≤−2.0 at total hip or spine (L1–L4) within 3 months prior to screening and clinical risk factors. Exclusion criteria: Patients who had received prior therapy with bisphosphonates, parathyroid hormone, strontium ranelate, raloxifene, calcitonin, high-dose corticosteroids, or hormone replacement within 6 months prior to randomisation; patients with a fracture within 6 months prior to randomisation, secondary osteoporosis, primary hyperparathyroidism, and presence of contraindications to study drugs were excluded. Other exclusion criteria included calculated creatinine clearance <35 mL/min; serum calcium >2.75 mmol/L, or <2.00 mmol/L; serum alkaline phosphatase higher than 2.5 times the upper limit of normal; any kind of jawbone disease or infection that may necessitate oral surgery during the course of the study and any tooth extractions during the last 3 months; or surgery of the jaw during the last 6 months before inclusion in the study. Patients with a history of invasive malignancy of any organ system within the past 5 years (excluding basal cell or squamous cell carcinoma of the skin) were also excluded. Mean age: 68 years Ethnicity: Caucasian (99%) Prevalent vertebral fracture: Not stated

Trial/study	Patient characteristics
RUTH trial[36]	Inclusion criteria: Postmenopausal women ≥55 year of age, ≥1 year postmenopausal, and had established CHD or were at high risk for CHD. Exclusion criteria: Not stated. Mean age: 68 years Ethnicity: White (84%) Prevalent vertebral fracture: Not stated
Silverman et al. 2008 (93)[37]	Inclusion criteria: Generally healthy women between the ages of 55 and 85 years were eligible for study inclusion if they were at least 2 years postmenopausal and had osteoporosis, defined as low BMD or radiographically confirmed vertebral fractures. Subjects without prevalent vertebral fracture were required to have lumbar spine or femoral neck BMD T-scores between −2.5 and −4.0 (inclusive), whereas subjects with prevalent vertebral fracture (at least one mild vertebral fracture) were required to have lumbar spine and femoral neck BMD T-scores not worse than −4.0. Exclusion criteria: Women were excluded if they had diseases that may affect bone metabolism, conditions that could interfere with bone mineral densitometry, pathologic vertebral fractures, vasomotor symptoms requiring treatment, or serious conditions such as endometrial hyperplasia or carcinoma, abnormal vaginal bleeding, malignancy within 10 years of the study, endocrine disorders requiring treatment, or untreated malabsorption disorders. Subjects with an active or history of deep vein thrombosis, pulmonary embolism, or retinal vein thrombosis were also excluded, as were subjects with elevated fasting total cholesterol or triglyceride levels (≥310 or ≥300 mg/dl, respectively). The use of androgens, systemic oestrogen (except estriol 2.0 mg/d), topical oestrogen (>3 times per week), progestogens, SERMs, bisphosphonates, calcitonin, parathyroid hormone, and cholecalciferol (>50,000 IU per week) was prohibited within 6 months of screening. Mean age: 66 years Ethnicity: White (87%) Prevalent vertebral fracture: 56%
VERO trial[38]	Inclusion criteria: Ambulatory post-menopausal women older than 45 years of age with a BMD T score less than or equal to –1.50 SDs at the femoral neck, total hip, or lumbar spine. Participants had to have radiographic evidence of at least two moderate (i.e., a reduction in vertebral body height of 26% to 40%) or one severe (more than 40% reduction) prevalent vertebral fragility fracture according to the classification of Genant and colleagues. Exclusion criteria: Patients with unresolved skeletal diseases other than osteoporosis, malignant tumours in the 5 years before screening, osteonecrosis of the jaw, previous atypical subtrochanteric femoral fractures, risk factors for osteosarcoma, GI disorders contraindicating risedronate, significantly impaired hepatic function, or a calculated creatinine clearance less than 30 mL/min using the Cockcroft–Gault equation. We also excluded patients who had undergone kyphoplasty or vertebroplasty at three or more levels before randomisation or

Trial/study	Patient characteristics
	within the 6 months before randomisation. Participants had to have normal baseline serum albumin-corrected calcium, parathyroid hormone, and free thyroxine concentrations, and 25-hydroxy-vitamin D concentration greater than 23 nmol/L. Mean age: 72 years Ethnicity: White (98%) Prevalent vertebral fracture: 100%
VERT MN trial (EU analysis)[39]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures. Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, and use of calcitonin, calcitriol or vitamin D supplements within 1-month, anabolic steroids, oestrogen, oestrogen-related drugs or progestogen within 3 months, or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months. Mean age: 71 years Ethnicity: Not stated (patients in the European analysis were all from Europe) Prevalent vertebral fracture: >50%
VERT-MN trial (AUS+EU analysis)[39]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures. Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, and use of calcitonin, calcitriol or vitamin D supplements within 1-month, anabolic steroids, oestrogen, oestrogen-related drugs or progestogen within 3 months, or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months. Mean age: 71 years Ethnicity: Not stated (patients were recruited in 80 European and Australian centres) Prevalent vertebral fracture: >50%
VERT-MN trial (NAm analysis)[40]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures or one vertebral fracture and low lumbar-spine (L1-L4) BMD (defined as <-0.83 g/cm^2 (Hologic instrument) or ≤0.94 g/cm^2 (Lunar instrument)). Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, or received drugs known to affect bone metabolism (e.g. calcitonin, calcitriol or cholecalciferol supplements within 1 month; anabolic steroids, oestrogen, oestrogen-related drugs or progestins within 3 months; or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months). Mean age: 69 years Ethnicity: Not stated (patients were recruited in 110 North American centres) Prevalent vertebral fracture: 80%

Trial/study	Patient characteristics
ZONE trial[41]	Inclusion criteria: Subjects were male and female Japanese patients aged between 65 and 89 years, and were ambulatory patients who had been diagnosed with primary osteoporosis based on the Diagnostic Criteria for Primary Osteoporosis of the Japanese Society for Bone and Mineral Research; patients who have fragility fractures caused by low BMD (young adult mean <80 %; T score <−1.7), with between one and four vertebral fractures from the fourth thoracic to the fourth lumbar vertebra (Th4 to L4). Exclusion criteria: Key exclusion criteria were a history of bisphosphonate use within 2 years prior to the study; serious complications including the heart, liver, or kidney disease; creatinine clearance <35.0 mL/min or urinary protein ≥2+; serum calcium <8.0 mg/dL or >11.0 mg/dL; and undergoing or planning to undergo an invasive dental procedure of the jawbone, such as tooth extraction, at the time informed consent was obtained. Mean age: 74 years Ethnicity: Japanese (100%) Prevalent vertebral fracture: 100%

AUS = Australia; BMD = bone mineral density; CHD = coronary heart disease; DXA = dual-energy x-ray absorptiometry; EU = European Union; FSH = follicle-stimulating hormone; GI = gastrointestinal; NAm = North America; NMA = network meta-analysis; SD = standard deviation; SERM = selective oestrogen receptor modulator; SI = Système international (d'unités), English: International System of Units; USA = United States of America

Additionally, the rates of fractures were presented in the CS at different time points for all comparator treatments, including placebo. As such, it is possible to compare the fracture rates across studies for placebo, which should be similar if the populations are similar. Across all fracture types and time points, the variability in fracture rates between studies included in the same NMA were large: for new vertebral fractures, the fracture rates varied between 1.4% and 40.0% at 12 months, 3.7% and 24.6% at 24 months and 4.1% and 25.7% at 36 months; for non-vertebral fractures, the fracture rates varied between 3.0% and 11.3% at 12 months, 3.0% and 11.3% at 24 months and 4.2% and 14.4% at 36 months; and for hip fractures, the fracture rates varied between 0.2% and 1.2% at 12 months, 0.2% and 8.7% at 24 months and 0.7% and 3.9% at 36 months. While the variation in fracture rates was largest for smaller studies, larger studies also had large variation: this is problematic as we would expect smaller studies to have more variable fracture rates than larger studies, which should have much closer fracture rates if the populations were similar. This is not necessarily indicative of potential effect modification, as, so long as fracture rates in a population in the absence of treatment are not effect modifiers, differences in the fracture rates do not by themselves indicate potential bias. However, very different fracture rates for placebo arms indicate large differences between populations, and some of these differences may be between effect modifiers, leading to potentially very large and undetectable biases.

3.3.4 Risk of bias assessment of the trials in the NMAs

The RoB assessments from the company for ARCH, FRAME and STRUCTURE are presented in Table 3.20, and for all other studies in the NMAs in Table 121 of Appendix D of the CS.[8] The ERG has checked the RoB assessments for the ARCH, FRAME and STRUCTURE trials and has no concerns about these assessments. The ERG did not assess the RoB for trials providing non-romosozumab evidence in the NMAs.

Table 3.20: Quality assessment for ARCH, FRAME and STRUCTURE

Trial number (acronym)	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
Was randomisation carried out appropriately?	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	No
Were there any unexpected imbalances in drop-outs between groups?	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No
Did the analysis include an intention-to-treat analysis?	Yes	Yes	Yes

Trial number (acronym)	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
If so, was this appropriate and were appropriate methods used to account for missing data?			
Based on CS, Table 120, Appendix D.[8] Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination) CRD = Centre for Reviews and Dissemination; CS = company submission			

3.4 Critique of the indirect comparison and/or multiple treatment comparison

In total, 11 NMAs were presented by the company, covering five distinct outcomes at three timepoints. As the BMD outcomes were not included in the CE model, we will restrict the critique of the indirect comparisons to the nine NMAs of fracture outcomes. We will also limit the critique to NMAs using the ITT populations, rather than the EU label populations. Furthermore, we will critique each of the NMAs separately, with reference to the population characteristics detailed in Table 3.19 above, which details the inclusion/exclusion criteria, mean age, ethnicity and prevalent vertebral fracture rate in each study. It is unclear whether age, ethnicity and prevalent vertebral fractures are effect modifiers, but in the view of the ERG, they are all plausible effect modifiers, and thus imbalances in these variables between trials may bias any analyses.

In general, apart from the potential biases from differences in effect modifiers, the ERG believes the NMAs to be well conducted.

3.4.1 New vertebral fractures

3.4.1.1 12 months

Figure 3.4 shows the network of evidence for the analysis of new vertebral fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo ([redacted])), although it is unclear if this is due to a lack of precision. The ERG asked the company to give both the direct and indirect results for all comparisons to judge whether the inconsistency estimates were imprecise or null and precise, but the company did not provide this information. However, it is still likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.4: Network of evidence for the analysis of new vertebral fractures at 12 months

Based on Figure 7 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the comparison. The FRAME study contributes the majority of evidence for any comparisons between romosozumab and comparator treatments except alendronate, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were different in all three trials. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: there was evidence of [redacted] heterogeneity for the comparison between raloxifene and placebo ([redacted]). The mean ages varied in all trials (65 to 74 years), the ethnicities varied (three trials were not international, conducted entirely within China, Japan or the USA, compared with ARCH and FRAME which were international), and the vertebral fracture rates were similar to FRAME but not ARCH. Therefore, there is a **very high risk of bias** from effect modification in the comparison between romosozumab and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Risedronate: the VERT-MN trials were conducted in Europe, Australia and North America, but had relatively similar characteristics to the ARCH trial, though higher prevalent vertebral fracture rates than FRAME. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab and risedronate.
- Teriparatide: the VERO trial compared teriparatide and risedronate, and therefore any RoB in the comparison between romosozumab and risedronate remains in this comparison, along with any RoB for the comparison between teriparatide and risedronate. The VERO trial included only

patients with prevalent vertebral fractures, and 98% of the patients were white, which is reasonably similar to the VERT-MN trials. There is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab and teriparatide.

3.4.1.2 24 months

Figure 3.5 shows the network of evidence for the analysis of new vertebral fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.5: Network of evidence for the analysis of new vertebral fractures at 24 months

Based on Figure 10 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of a RoB from effect modification between the ARCH and FIT I trials, although the FIT I trial did not report the ethnicity of patients, and there may be effect modification from unmeasured variables. The company stated that there was no evidence of inconsistency for the two closed-loops in the network (risedronate – placebo – teriparatide (P=[redacted]), and teriparatide – placebo – abaloparatide (P=[redacted])).

- Zoledronate: There was [redacted] evidence for heterogeneity for the comparison between zoledronate and placebo (I^2 = [redacted]). However, the patient ethnicities were different between these trials (one study was conducted solely in China, one in Japan, and one was international), the mean ages of patients was markedly different (between 57 years and 74 years), though the rate of prevalent vertebral fractures was high in all studies, as in ARCH. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.

- Raloxifene: The MORE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not report the ethnicity or location of patients. Therefore, there is **a high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: the VERT-MN AUS trial was conducted in Australia, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.
- Teriparatide: there was evidence of [redacted] heterogeneity for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.1.3 36 months

Figure 3.6 shows the network of evidence for the analysis of new vertebral fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.6: Network of evidence for the analysis of new vertebral fractures at 36 months

Based on Figure 13 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of effect modification between the ARCH and FIT I trial, though the mean age of patients and rate of prevalent vertebral fractures were both lower in the Liberman 1995 trial, and the FIT I trial did not report the ethnicity of patients. There was [redacted] observed heterogeneity between the comparison of alendronate and placebo (I^2 = [redacted]).

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: There was [redacted] observed heterogeneity between the comparison of raloxifene and placebo (I^2 = [redacted]). The MORE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not report the ethnicity or location of patients. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Australia and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.2 Non-vertebral fractures

3.4.2.1 12 months

Figure 3.7 shows the network of evidence for the analysis of non-vertebral fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo ([redacted])), although it is unclear if this is due to a lack of precision and is close to statistical significance. It is likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.7: Network of evidence for the analysis of non-vertebral fractures at 12 months

Based on Figure 16 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. The FRAME study will contribute the majority of evidence for any comparisons between romosozumab and comparator treatments, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: there was evidence of [redacted] heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all 3 trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Risedronate: the VERT-MN EU trial was conducted in Europe, but had relatively similar characteristics to the ARCH trial, though higher prevalent vertebral fracture rates than FRAME. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab and risedronate.
- Teriparatide: The Neer 2001 trial did not include any Hispanic patients and only included patients with prevalent vertebral fracture. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab and teriparatide.

3.4.2.2 24 months

Figure 3.8 shows the network of evidence for the analysis of non-vertebral fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.8: Network of evidence for the analysis of non-vertebral fractures at 24 months

Based on Figure 19 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is some evidence of a risk of bias from effect modification between the ARCH and FOSIT trials, as the FOSIT trial included younger patients than the ARCH trial (mean of 63 years vs. 74 years), and the FOSIT trial did not report the rate of prevalent vertebral fractures. The company stated that there was [redacted] evidence of inconsistency for the closed-loop in the network (risedronate – placebo – teriparatide (P=[redacted])).

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the ARCH trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.

- Risedronate: the VERT-MN EU trial was conducted in Europe, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.
- Teriparatide: there was [redacted] observed heterogeneity for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.2.3 36 months

Figure 3.9 shows the network of evidence for the analysis of non-vertebral fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. As there was no data for non-vertebral fractures at 36 months in the ARCH trial, data from 30 months was used instead, which may have caused bias. It is possible, therefore, that the comparison between romosozumab/alendronate and alendronate has some bias, which will propagate to all other comparisons.

Figure 3.9: Network of evidence for the analysis of non-vertebral fractures at 36 months

Based on Figure 22 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is, however, little evidence of a RoB from effect modification between the ARCH and FIT I+II trials, though the FIT trials did not report the ethnicity of patients.

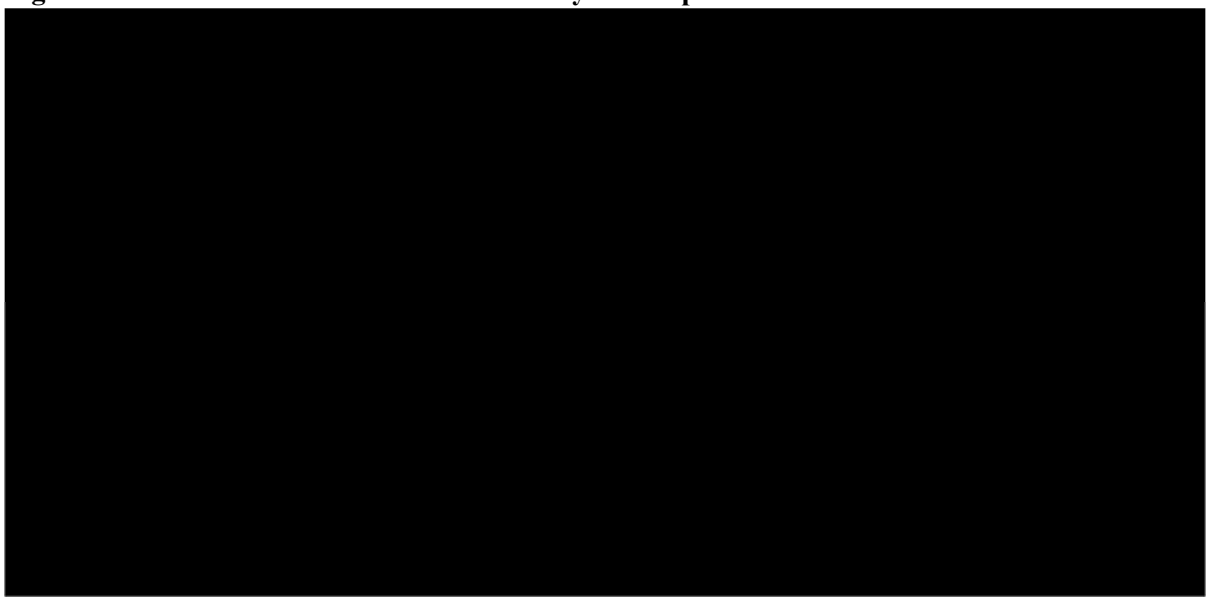
- Zoledronate: There was [redacted] observed heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: there was [redacted] observed heterogeneity between the comparison of raloxifene and placebo (I^2 = [redacted]). The RUTH was relatively similar to the ARCH trial. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: there was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Europe and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.3 Hip fractures

3.4.3.1 12 months

Figure 3.10 shows the network of evidence for the analysis of hip fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo [[redacted]]), although it is unclear if this is due to a lack of precision. It is therefore likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.10: Network of evidence for the analysis of hip fractures at 12 months



Based on Figure 25 of Appendix D of the CS.[8]
CS = company submission

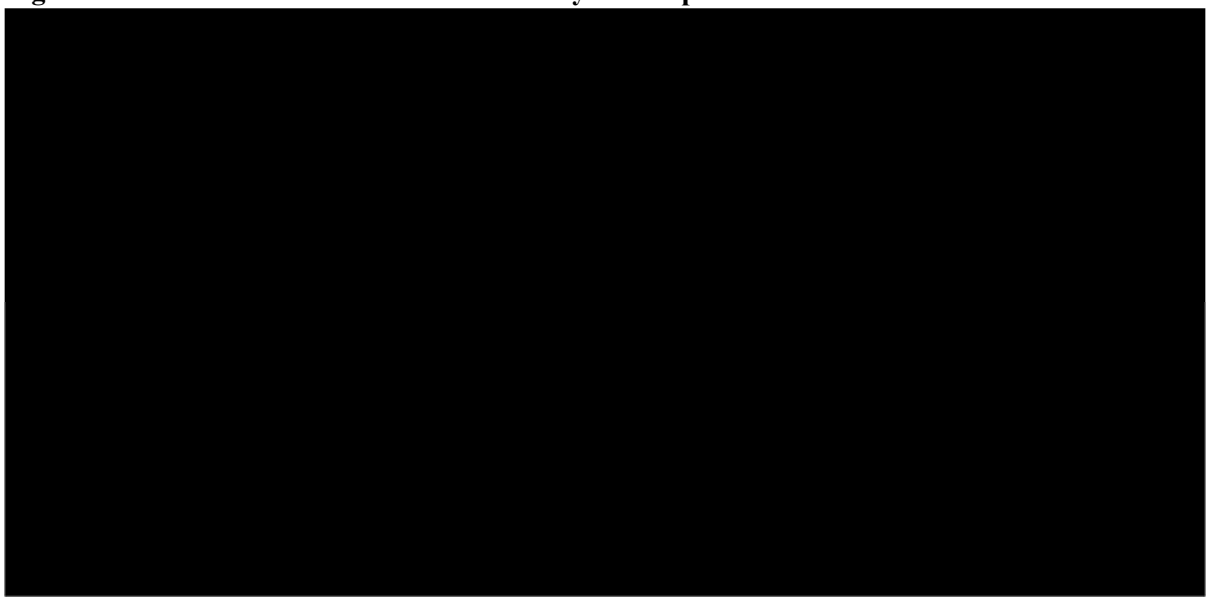
All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. The FRAME study will contribute the majority of evidence for any comparisons between romosozumab and comparator treatments, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: there was [redacted] observed heterogeneity for the comparison of zoledronate and placebo (I^2 = [redacted]). Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all three trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the FRAME trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.

3.4.3.2 24 months

Figure 3.11 shows the network of evidence for the analysis of hip fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB as the ARCH trial does not have a high RoB.

Figure 3.11: Network of evidence for the analysis of hip fractures at 24 months



Based on Figure 28 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of effect modification between the ARCH and FIT I trials, although the FIT I trial did not report the ethnicity of patients, and there may be effect modification from unmeasured variables. The company stated that there was [redacted] evidence of inconsistency for the closed-loop in the network (teriparatide – placebo – abaloparatide [P=[redacted]]).

- Zoledronate: There was [redacted] evidence of heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. The Bai 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the ARCH trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was evidence of [redacted] heterogeneity between the comparison of teriparatide and risedronate (I^2 = [redacted]). The VERO trial compared teriparatide and risedronate, and therefore any RoB in the comparison between romosozumab and teriparatide remains in this comparison, along with any RoB for the comparison between teriparatide and risedronate. In the VERO trial, 98% of the patients were white, while in ARCH 32% of patients were Hispanic and 68% of patients were non-Hispanic. The Hadji 2012 trial had similar patient

characteristics as the ARCH trial. There is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.

- Teriparatide: the company state that [redacted] heterogeneity was observed for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.3.3 36 months

Figure 3.12 shows the network of evidence for the analysis of hip fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. As there was no data for hip fractures at 36 months in the ARCH trial, data from 30 months was used instead, which may have caused bias. It is possible, therefore, that the comparison between romosozumab/alendronate and alendronate has some bias, which will propagate to all other comparisons.

Figure 3.12: Network of evidence for the analysis of hip fractures at 36 months

Based on Figure 31 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (since the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence for a RoB from effect modification between the ARCH and FIT I trials, and the FIT I trial did not report the ethnicity of patients.

- Zoledronate: There was [redacted] observed heterogeneity between the comparison of alendronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH and FRAME, but

patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all 3 trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.

- Raloxifene: The RUTH was relatively similar to the ARCH trial. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Europe and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.4 Summary

Overall, there was little direct evidence for comparisons for romosozumab included in any of the NMAs, and most studies had differences in mean age, ethnicity or rate of prevalent vertebral fractures, indicating at least a moderate RoB from effect modification. Additionally, as almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone.

Additionally, individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the RoB. As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low RoB; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high RoB.

3.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work on clinical effectiveness was undertaken by the ERG.

3.6 Conclusions of the clinical effectiveness section

The decision problem is largely in line with the NICE scope. However, the population in the CS is postmenopausal women with severe osteoporosis at high risk of fracture; where "high risk of fracture" is defined as having suffered a fracture within the last 2 years (also revert to as "imminent risk of fracture").[1] This is narrower than the population in the NICE scope (Postmenopausal women with severe osteoporosis at high risk of fracture; where timing of previous fracture is not mentioned),[2] and

narrower than the population in the ARCH trial (Postmenopausal women with severe osteoporosis at high risk of fracture; where timing of fracture is not an inclusion criterion for some patients).[3]

The clinical effectiveness evidence for romosozumab in severe osteoporosis in the CS is mainly based on the ARCH trial. Two other phase III clinical trials, the FRAME and STRUCTURE trials are mentioned in the CS as well. However, neither the FRAME nor STRUCTURE trials studied a patient population aligned to where the company expects romosozumab to be used in National Health Service (NHS) clinical practice.[17, 18] A fourth study, the BRIDGE study, considered use in men, which is not part of the marketing authorisation for romosozumab.[19]

The ARCH trial is a phase III, multicentre, randomised, double-blind trial, comparing romosozumab followed by alendronate vs. alendronate alone in postmenopausal women with severe osteoporosis and a fragility fracture (see Table 3.5).[3] This trial provides evidence for romosozumab in its expected position in the clinical pathway: a first-line therapy in patients who have previously suffered a MOF. Efficacy outcomes reported in ARCH include incidence of clinical, vertebral, non-vertebral and hip fracture and percentage change from baseline in BMD. Data from ARCH were used as the main data for the economic modelling in this submission.

In the ARCH trial, romosozumab/alendronate statistically significantly reduced the incidence of new vertebral fractures at month 24, meeting its primary endpoint. Patients in the romosozumab/alendronate arm had a 50% lower relative risk of vertebral fractures compared to patients on alendronate alone over 24 months (RR 0.50, 95% CI 0.38 to 0.66).[20] Additionally, a statistically significantly lower proportion of patients experienced a clinical fracture (non-vertebral fracture and clinical vertebral fracture) at the time of primary analysis in the romosozumab/alendronate group compared to alendronate alone (HR 0.73, 95% CI 0.61 to 0.88), meeting the other primary endpoint.[20] At the primary analysis there were also a lower number of patients who experienced non-vertebral fractures (HR 0.81, 95% CI 0.66 to 0.99) and hip fractures (HR 0.62, 95% CI 0.42 to 0.92).[20] Patients treated with romosozumab also had a statistically significantly greater increase in BMD from baseline compared to alendronate (adjusted $P<0.001$), which was maintained until month 36.[20]

The Kaplan-Meier curves for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3) show that there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first fracture up to month 42. However, the curves seem to converge again by month 48. This means that it is possible that the effects of romosozumab wane over time. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Therefore, longer term follow-up is needed to see whether the effects are maintained over time.

Overall, results of the ARCH trial are favourable for romosozumab. Both primary outcomes (the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis) are met and most fracture results significantly favour romosozumab over alendronate. In addition, all BMD outcomes significantly favour romosozumab over alendronate. However, the graphs for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3), seem to indicate that the effectiveness of romosozumab over alendronate becomes less after 42 months; longer term follow-up is needed to see whether the effects are maintained over time.

The incidences of AEs and SAEs were similar overall in the ARCH trial between the two treatment groups during the 12-month double-blind period, and cumulative incidences were similar between the two groups during the primary analysis period. However, more people in the romosozumab group

experienced adjudicated serious CV AEs during the double-blind period, with 50 patients (2.5%) in the romosozumab group and 38 (1.9%) in the alendronate group reporting these events (odds ratio (OR) 1.31, 95% CI 0.85 to 2.00). A total of 16 patients (0.8%) in the romosozumab group and 6 (0.3%) in the alendronate group reported cardiac ischemic events (OR 2.65; 95% CI, 1.03 to 6.77), and 16 patients (0.8%) in the romosozumab group and 7 (0.3%) in the alendronate group reported cerebrovascular events (OR 2.27, 95% CI 0.93 to 5.22). Therefore, romosozumab is contraindicated for patients with a history of MI or stroke.

The company claims that it is *"reasonable to conclude that a population of patients treated with romosozumab will experience a reduced level of pain, disability and mortality, relative to patients treated with currently available treatments, because these patients will experience fewer fragility fractures compared to patients treated with currently available treatments"* (Response to request for clarification, question A11).[9] However, after 12 months, more patients died in the romosozumab group (n=30, 1,5%) than in the alendronate group (n=21, 1.0%). At the time of the primary analysis, 90 patients had died in both groups.

In total, 11 NMAs were presented by the company, covering five distinct outcomes at three timepoints, although only the three fracture outcomes were used in the CE model. In these NMAs, many comparator treatments were directly and indirectly compared with romosozumab using Bayesian methods. The methods used appear valid and appropriate.

However, there was little direct evidence for comparisons for romosozumab included in any of the NMAs, and most studies had differences in mean age, ethnicity or rate of prevalent vertebral fractures. As these variables could potentially be effect modifiers when conducting indirect comparisons, different levels of these variables in the included studies likely indicates at least a moderate risk of bias from effect modification. Additionally, as almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone.

Additionally, individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the RoB. As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low RoB; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high RoB.

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

One set of systematic literature searches was performed to identify CE studies and costs and healthcare resource use studies (CS Appendix G and Appendix I).[8] Searches were not conducted to identify health-state utility values. Instead, economic evaluations included in the original and update economic systematic literature reviews (SLRs) were reviewed for novel health-state utility values of relevance to the CE model for romosozumab.

4.1.1 Searches performed for cost effectiveness section

Appendices G and I of the CS reported the literature searches used to identify CE studies and costs and healthcare resource use studies.[8] Searches were conducted in March and April 2018 and an update search was conducted in February and March 2021. Summaries of the resources searched are provided in Tables 4.1 and 4.2. The following paragraphs contain summaries and critiques of all searches related to CE presented in the CS.

Table 4.1: Resources searched for cost effectiveness studies and costs and healthcare resource use studies. March/April 2018

Resource	Host/Source	Date Range	Date searched
Databases			
Embase	OvidSP	1974 to 9th March 2018	9 March 2018
MEDLINE, including MEDLINE Daily, MEDLINE In-Process and Epub Ahead of Print	OvidSP	1946 to 9th March 2018	9 March 2018
NHS EED	Cochrane Library: Wiley Online	Issue 2 of 4, April 2015	9 March 2018
HTA Database	Cochrane Library: Wiley Online	Issue 4 of 4, October 2016	9 March 2018
EconLit	EBSCO	1886 to 8th March 2018	9 March 2018
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear4/	-	13 April 2018
ScHARRHUD	www.scharrhud.org/	-	13 April 2018

Resource	Host/Source	Date Range	Date searched
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/	-	13 April 2018
Conference Proceedings			
WCO-IOF-ESCEO	PDF abstract books	2016 and 2017	13 April 2018
ECTS	PDF abstract books	2016 and 2017	13 April 2018
ASBMR	PDF abstract books	2016 and 2017	13 April 2018
ISPOR	https://www.ispor.org/heor-resources/presentations-database/search	2016 and 2017	13 April 2018
FFN	PDF abstract book	2016 and 2017	13 April 2018
EULAR	http://scientific.sparx-ip.net/archiveeular/	2016 and 2017	13 April 2018
HTA websites			
NICE	https://www.nice.org.uk/	-	13 April 2018
SMC	https://www.scottishmedicines.org.uk/	-	13 April 2018
AWMSG	www.awmsg.org/	-	13 April 2018
NCPE	http://www.ncpe.ie/	-	13 April 2018

The bibliographies of all relevant SLRs, meta-analyses and HTA submissions identified through the electronic database and HTA agency website searches were also manually searched to identify any additional studies of relevance.

ASBMR = American Society for Bone and Mineral Research; AWMSG = All Wales Medicines Strategy Group; CEA = cost effectiveness analysis; ECTS = European Calcified Tissue Society; EED = Economic Evaluation Database; EQ-5D = EuroQol-5 Dimensions; EULAR = European League Against Rheumatism; FFN = Fragility Fracture Network; HTA = health technology assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NCPE = National Centre for Pharmacoeconomics; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; ScHARRHUD = School of Health and Related Research Health Utilities Database; SLR = systematic literature review; SMC = Scottish Medicines Consortium; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases

Table 4.2: Resources searched for cost effectiveness studies and costs and healthcare resource use studies. February/March 2021

Resource	Host/Source	Date Range	Date searched
Databases			
Embase	OvidSP	1974 to 24th February 2021	24 February 2021
MEDLINE, including MEDLINE Daily, MEDLINE In-Process and Epub Ahead of Print	OvidSP	1946 to 24th February 2021	24 February 2021
NHS EED	Centre for Reviews and Dissemination	Issue 2 of 4, April 2015	24 February 2021
HTA Database	Centre for Reviews and Dissemination	Issue 4 of 4, October 2016	24 February 2021
INAHTA HTA Database	Not reported	from 1996 to 24th February 2021	24 February 2021
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear4/	-	5 March 2021
ScHARRHUD	www.scharrhud.org/	-	5 March 2021
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/	-	5 March 2021
Conference Proceedings			
WCO-IOF-ESCEO	PDF abstract books	2019 and 2020	5 March 2021
ECTS	PDF abstract books	2019 and 2020	5 March 2021
ASBMR	PDF abstract books	2019 and 2020	5 March 2021
ISPOR	https://www.ispor.org/heor-resources/presentations-database/search	2019 and 2020	5 March 2021
FFN	PDF abstract book	2019	5 March 2021
EULAR	http://scientific.sparx-ip.net/archiveeular/	2019 and 2020	5 March 2021

Resource	Host/Source	Date Range	Date searched
HTA websites			
NICE	https://www.nice.org.uk/	-	5 March 2021
SMC	https://www.scottishmedicines.org.uk/	-	5 March 2021
AWMSG	www.awmsg.org/	-	5 March 2021
NCPE	http://www.ncpe.ie/	-	5 March 2021
The bibliographies of all SLR or (network) meta-analyses ([N]MAs) identified in the course of this update were hand-searched in order to identify any additional, relevant studies for inclusion. ASBMR = American Society for Bone and Mineral Research; AWMSG = All Wales Medicines Strategy Group; CEA = cost effectiveness analysis; ECTS = European Calcified Tissue Society; EED = Economic Evaluation Database; EQ-5D = EuroQol-5 Dimensions; EULAR = European League Against Rheumatism; FFN = Fragility Fracture Network; HTA = health technology assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NCPE = National Centre for Pharmacoeconomics; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; ScHARRHUD = School of Health and Related Research Health Utilities Database; SLR = systematic literature review; SMC = Scottish Medicines Consortium; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases			

ERG comment:

- The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported.
- Conference proceedings were searched. Full details of the conference searches, including search terms, URLs, results and the date of the searches, were provided. A full explanation for the two-year date limit was provided.
- Additional health economic specific resources were searched, and full details of the search strategies or search terms used, dates of searches, and results, were reported in the CS.[1, 8]
- Health technology assessment organisation websites were searched, and full details of the search terms used, dates of searches, and results, were reported in the CS.[1, 8]
- Extensive use of truncation, proximity operators, synonyms, and subject headings (MeSH and EMTREE) were included in the search strategies. Study design search filters for CE evaluations and UK cost studies were included. It would have been helpful if the search filters had been cited in the methods section.[72] There were no language or date limits.
- Update searches were conducted in February and March 2021. Full details of the searches were provided.
- Searches of NHS EED and the HTA database for the original review searches were conducted via the Cochrane Library. These resources were no longer available via the Cochrane Library by the time of the update searches in February 2021, so the CS translated the searches to run in the Centre for Reviews and Dissemination (CRD) interface. In addition, the company searched the International Network of Agencies for Health Technology Assessment (INAHTA) Database to retrieve more up-to-date health technology assessment reports. A full explanation for these changes was provided in the CS.[1, 8]
- No searches were conducted to identify health-state utility values. The CS reported that *"To supplement the search for economic data, all economic evaluations included in the original and*

update SLRs were reviewed for novel health-state utility values of relevance to the cost-effectiveness model for romosozumab. The economic evaluations were reviewed by two independent reviewers and their results compared to reach consensus. Any disagreements were resolved by a third independent reviewer, if necessary." The company did search health utilities resources (CEA Registry, ScHARRHUD and EQ-5D Publications Database).

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on CE studies, utilities and costs and resource use are presented in Table 4.3.

Table 4.3: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	Men and/or postmenopausal women with osteoporosis at increased risk of fracture. Patients may be stated to be at 'risk of fracture' in the paper, or may have been defined as at risk by the presence of at least one of the following: • Age ≥65 years (women) and ≥75 years (men) • BMD T-score of ≤2.5 • Prior fracture • Family history • Long periods of inactivity	• Patients being studied for the prevention or treatment of glucocorticoid-induced osteoporosis • Patients with normal or unspecified BMD who have not been selected based on the presence of risk factors (see left) • Patients with other indications for osteoporosis treatment, including: • Hormonal disorders, e.g., hyperthyroidism, pituitary gland disorders, Cushing's syndrome, hypogonadism • Paget's disease • Hypercalcaemia of malignancy • Breast cancer • Prostate cancer • Rheumatoid arthritis • Coeliac disease • Crohn's disease • Eating disorders, e.g., bulimia or anorexia • Heavy smoking or drinking Where studies included a mixed population of participants in which the above eligibility criteria were not met by all patients, the study was excluded unless separate data on the outcomes of interest were reported for the population of interest.
Intervention (economic evaluations)	Romosozumab, or any of the below interventions: • Teriparatide • Bisphosphonates (alendronate, risedronate, ibandronate and zoledronic acid)	• Combination therapies (with the exception of combination of an intervention of interest with vitamin D and calcium supplementation)

	Inclusion criteria	Exclusion criteria
	• Denosumab • Raloxifene • Strontium ranelateb • Abaloparatideb	• Interventions of interest that were co-administered with any other therapy, with the potential to augment bone, unless concomitant treatments were specified in the summary of product characteristics • Interventions that were not administered in accordance with their licensed indication
Intervention (cost and resource use)	Any or none	Not applicable.
Comparator	Any or no comparator	Not applicable.
Outcomes(s) 1 (Published economic evaluations)	Outcomes of relevant study designs, including: • Costs, including cost per fracture event avoided • Life years gained • Quality-adjusted life years • Number of fractures • Number of patients with fractures • Incremental costs and QALYs • Incremental cost effectiveness ratios	Studies not reporting relevant outcomes
Outcomes(s) 3 (Cost/resource use studies)	Original direct costs or resource use data published in 2008 onwards relevant to an economic model of romosozumab in the prevention of fractures in osteoporosis, including but not necessarily limited to: • Treatment and management of fractures, including: • Fractures of the hip and vertebrae • Nursing home/long-term care • BMD measurement • Physician visits • Proton pump inhibitor for gastrointestinal events • IV injections of zoledronate and denosumab • Nurse visit • Distribution of patients among treatment sites, including: • Hospital (inpatient and outpatient) • Accident and emergency department • Nursing home Data must be relevant to the UK NHS and Personal and Social Services	Studies not reporting relevant outcomes, or reporting indirect costs only

	Inclusion criteria	Exclusion criteria
Study design 1 (Economic evaluations)	Original economic evaluations considering both the costs and benefits of alternative interventions. Specifically, the following types of analysis: • Cost effectiveness • Cost utility • Cost benefit • Cost minimisation • Cost consequence To be eligible, models needed to be novel with a base-case in the UK, US, Australia or Canada. Non-novel models were only eligible if the base-case was the UK.	• Publications without original data • Study protocol reporting no results • Comments • Letters • Editorials • Non-systematic/narrative reviews • Animal/in vitro studies
Study design 3 (Cost/resource use studies)	Primary research publications on any study design	• Publications without original data • Study protocol reporting no results • Comments • Letters • Editorials • Non-systematic/narrative reviews • Animal/in vitro studies
Publication type (economic evaluations)	• Journal articles presenting original research • HTAs presenting primary research • Original SLR: Congress abstracts published in or after 2016 • During SLR update: Congress abstracts published in or after 2019	Other publications types
Publication type (cost and resource use)	• Journal articles presenting original research • SLRs of relevant primary publications (these were included at the title/abstract review stage and were used for the identification of any additional primary studies not identified through the database searches. They were excluded during the full-text review unless they reported primary, original research themselves) • HTAs presenting primary research • Original SLR: Congress abstracts published in or after 2016 • During SLR update: Congress abstracts published in or after 2019	

	Inclusion criteria	Exclusion criteria
Other (Economic evaluations)	• English language only • Human subjects only	• Articles not in the English language • Studies not in human subjects
Other (cost and resource use)	• Studies conducted in the UK • English language only • Human subjects only	• Articles not in the English language • Studies not conducted in the UK • Studies not in human subjects

Based on Tables 142 and 156 from the Appendices of the CS.[8]
[a] If a study did not specifically state that women were postmenopausal, then it was not excluded. However, if a study specifically stated that patients were not postmenopausal, it was excluded; [b] Strontium ranelate and teriparatide were included as potentially relevant comparators at the time of the original SLR, which was conducted before the NICE Scope was released.
BMD = bone mineral density; CS = company submission; HTA = health technology assessment; ICER = incremental cost effectiveness ratio; IV = intravenous; LYG = life years gained; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal and Social Services; QALY = quality-adjusted life year; SLR = systematic literature review; UK = United Kingdom; US = United States

In total, 3,732 unique articles were reviewed at the title/abstract review stage in the economic evaluation SLR.[8] Of these, 352 articles were deemed potentially relevant and reviewed at the full-text stage, with 29 articles ultimately meeting the economic evaluation inclusion criteria and three meeting the cost/resource use criteria. An additional nine articles were identified through congress searching, website searching and through handsearching of bibliographies in the economic evaluation SLR, resulting in a total of 38 articles reporting on 35 unique studies being included. These studies are summarised in Tables 143 and 147 of the CS appendices.[8] No additional cost and resource use articles were identified, resulting in a total of three studies being included in this review. These studies are summarised in Table 157 of the CS appendices.[8]

An additional SLR for HRQoL was not conducted. All economic evaluations included in the original and updated SLRs were reviewed for novel health-state utility values of relevance to the CE model for romosozumab. The handsearching of included economic evaluations did not identify any novel health-state utility values of relevance to the romosozumab model.

4.1.3 Conclusions of the cost effectiveness review

The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported. Overall, the ERG does not have any major concerns regarding the searches but notes that no searches were conducted to identify health-state utility values (see Section 4.1.1 for more details), it is unclear whether empirical studies estimating utility values in this condition were missed as only included economic evaluations were searched for utility values. Furthermore, it is unclear whether relevant resource use data were missed by including only studies conducted in the UK. Resource use data from other countries could have been considered, with UK unit costs applied.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.4: NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers.	As per the reference case.
Perspective on costs	NHS and PSS.	As per the reference case.
Type of economic evaluation	Cost utility analysis with full incremental analysis.	As per the reference case.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared.	As per the reference case.
Synthesis of evidence on health effects	Based on systematic review.	As per the reference case.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	As per the reference case.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers.	Utility multipliers for fracture events were estimated from patient reported data from the ICUROS study. These multipliers were applied to UK general population EQ-5D norms.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population.	Although not explicitly stated, it seems that the UK EQ-5D valuation tariff has been used to estimate the multipliers. The UK value set was used to the was used to estimate the general population norms.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	As per the reference case.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS.	As per the reference case.
Discounting	The same annual rate for both costs and health effects (currently 3.5%).	As per the reference case.

Element of health technology assessment	Reference case	ERG comment on company's submission
ERG = Evidence Review Group; HRQoL = health related quality of life; ICUROS = International Costs and Utilities Related to Osteoporotic Fractures Study; NHS = National Health Service; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom		

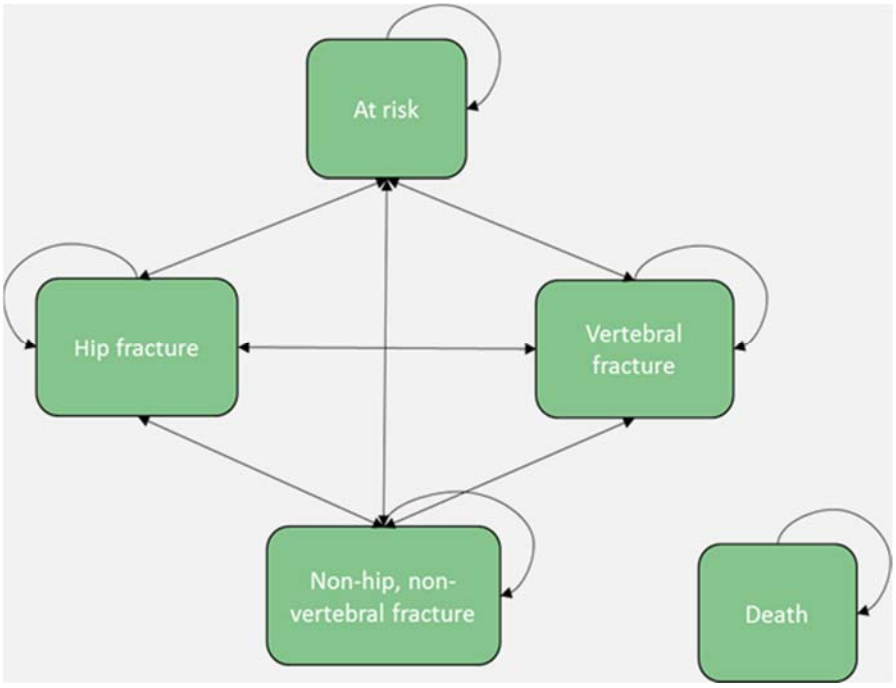
4.2.2 Model structure

4.2.2.1 Health states/events and transitions

A "*de novo*" Markov microsimulation model was developed in Microsoft Excel to assess the CE of romosozumab followed by alendronate compared to alendronate alone in postmenopausal women who have experienced a MOF within the past 24 months.

The model, shown in Figure 4.1, consisted of five health states: at risk, hip fracture, vertebral fracture, NHNV fracture and death.

Figure 4.1: Model structure



Based on Figure 13 in the CS.[1]
CS = company submission

At the start of the model, all patients are in the "at risk" health state. At the end of each model cycle patients can either transition to one of the fracture states, stay in the same health state without having a new fracture, or die. Upon transitioning to "death", patients remain there for the rest of the simulation. No restrictions were imposed for the sequence or number of fractures experienced.

As an advantage of the micro-simulation approach, the model keeps track of each patient's history to enable the calculation of costs, quality of life, and fracture risk over the lifetime (with a maximum of 100 years) of each individual patient.

At any point in the model, the risk of sustaining a fracture is based on a combination of four components:

1. The general population risk of fracture.

2. The increased fracture risk associated with osteoporosis, relative to the general population.
3. The increased fracture risk due to having sustained a recent fracture (i.e., the imminent fracture risk).
4. The reduction in risk, where applicable, due to osteoporosis treatment.

The input values, and their underlying assumptions, for each of these components are further elaborated in Section 4.2.6 of the ERG report.

The same model, but with different input values, was also used as the basis for two recent publications in the peer-reviewed journal of the International Osteoporosis Foundation: 'Osteoporosis International'.[73, 74] In Söreskog et al. 2021a the CE of romosozumab followed by alendronate compared to alendronate alone for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture was assessed from a Swedish perspective with model inputs for treatment effectiveness based on ARCH.[73] In Söreskog et al. 2021b the CE of a ("hypothetical") bone-forming agent followed by an anti-resorptive therapy compared to an anti-resorptive therapy alone was assessed for the prevention of fractures in patients with osteoporosis from a UK perspective.[74]

ERG comment: The model structure appears appropriate. However, the ERG's ability to step through and evaluate the model functionality was hindered by the fact that all model calculations are done in background VBA code. The VBA code is password protected and the company were unable to make the password available to the ERG due to confidentiality issues with the FRAX algorithm that was implemented in the VBA code. Outside of the VBA code only input parameters and hardcoded results are available. At clarification, the company did provide some of the VBA code in separate files but the ERG was unable to:

- Verify that this matched the code within the model.
- Step through the code as they would in the model to understand the functionality of the code.
- Make any changes to the code in response to potential errors or to make ERG or base-case changes (beyond changes to the available input parameters).

At a later stage a version of the model was made available to the ERG in which the VBA code was separated in a non-password protected version for the code that was not related to FRAX and a version with the password protected FRAX algorithm. However, the ERG was advised not to use this version of the model for running analyses. Therefore, the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. This means that the ERG has not been able to carry out its usual level of investigation and has had to proceed by assuming that the model functions correctly and as reported by the company.

The ERG comment in Section 5.3 presents some inconsistencies and issues found in the model and the VBA code. These appear to have a minor impact on the results, but this needs further confirmation from the company.

4.2.3 Population

The population in the Final Scope by NICE is defined as "Postmenopausal women with severe osteoporosis at high risk of fracture", in line with the marketing authorisation by the European Medicines Agency (EMA) for the use of romosozumab in women who have been through the menopause and who have severe osteoporosis (low bone density and previous fracture), leading to a high risk of further fractures. Severe osteoporosis is defined, according to the World Health Organization (WHO), based on a BMD value below a T-score of −2.5 and with one or more fragility

fractures (i.e., low impact fractures sustained from standing height or less). Importantly, the NICE final scope does not define "high risk of fracture".

The modelled population in the CS consisted of postmenopausal women with baseline characteristics, provided in Table 4.5, that are the average of those in the trial population in ARCH in terms of age (i.e., 74 years), femoral neck BMD T-score (i.e., -2.90) and BMI (i.e., 25.41). The inclusion criteria used in ARCH are listed in Section 3.2.1 of the ERG report. As described in that Section as well, the modelled population in the CS is assumed to consist of patients who have had a MOF within the prior 24 months. Based on the FRAX algorithm in combination with the additional risk that is associated with a recent fracture, the modelled population had an estimated mean 10-year MOF probability of 30%. An important difference between the ARCH ITT population and the modelled population is that ARCH included patients who previously sustained a fracture regardless of recency, whereas for the modelled population it is assumed that a previous fracture was sustained within 24 months prior to the start of treatment. In the ARCH ITT population, [redacted] of patients suffered a MOF within 24 months prior to randomisation.

Table 4.5: Baseline patient characteristics used in the economic model

Model parameter	Value	Source and appropriateness for modelling patient population in decision problem
Sex	Female	Licensed indication
Fracture history	Recent fracture (MOF within 24 months)	ARCH,[3] Swedish registry.[75] Specifying MOF aligns with the expected target population for romosozumab in clinical practice, to maximise the benefits of treatment
Mean age, years	74	ARCH[3]; comparable to the average age of postmenopausal women with osteoporosis in the UK [11, 76]
Mean femoral neck T-score (SD)	−2.90	ARCH[3]
Mean BMI	25.41	ARCH[3]
Mean 10-year MOF probability	30%	Target patient population

Based on Table 17 of the CS.[1]
BMI = body mass index; CS = company submission; MOF = major osteoporotic fracture; SD = standard deviation; UK = United Kingdom.

ERG comment: The issues with the population explained in Sections 2.1 and 3.2.1 are also applicable to the CE analyses.

4.2.4 Interventions and comparators

The modelled intervention consisted of a once-in-a-lifetime, 12-month course of romosozumab, followed by a 48-month course of alendronate. Romosozumab is administered monthly at a dose of 210 mg via two subcutaneous injections of 105 mg each into the abdomen, thigh, or upper arm. Alendronate is administered orally at a weekly dose of 70 mg.

The comparators that were used in the company base-case model consists of a 60-month course of alendronate, administered orally at a weekly dose of 70 mg, and no treatment. Additionally, the company performed a series of scenario analyses for which the following comparators were used i.e., instead of alendronate: teriparatide, denosumab, risedronate, zoledronate, and raloxifene. Teriparatide

is administered daily at a dose of 20 µg (i.e., microgram) via subcutaneous injection into the abdomen or thigh, over the course of (maximally) 24 months per lifetime. Denosumab is administered once every 6 months at a dose of 60 mg via a single subcutaneous injection into the thigh, abdomen or upper arm. Risedronate is administered orally once per week at a dose of 35 mg. Zoledronate is administered once per year via intravenous infusion at a dose of 5 mg. Raloxifene is administered orally at a daily dose of 60 mg. For all modelled comparators in the scenario analyses a treatment duration of 60 months was assumed, except teriparatide for which the maximum treatment duration of 24 months was assumed. A description of all the included treatment sequences, their durations and their residual effects is provided in Table 4.14 in Section 4.2.6.3.

ERG comment: The treatments that were used as comparators in the company's base-case and scenario analyses include all that were listed in the NICE scope, except for ibandronic acid. The company indicated that no trials for ibandronate at the licensed dose were found to be included in the NMA for fracture outcomes, and therefore this comparator was not included.

The ERG notes that there is uncertainty regarding the appropriateness and relevance of the included comparators, due to the uncertainty regarding the relevant population as described in the previous section. This is because risk of fracture is often used to guide choice of treatment.

For the information summarised above, the ERG noted some small inconsistencies in the information that was provided in Table 31 of the CS relative to information provided in the corresponding summaries of product characteristics and other general sources regarding medicines that can be found online.[1] Specifically, the ERG added the daily dose of teriparatide and corrected the dosage of zoledronate (5 mg instead of 4 mg) and frequency of administration for denosumab (once every 6 months instead of once every 6 weeks).

4.2.5 Perspective, time horizon and discounting

The analysis was performed from a NHS and Personal Social Services (PSS) perspective, in line with the NICE reference case.[77] The model used a lifetime time horizon, following a patient until either death or an age of 100 years, which was in line with both the NICE reference case and European Society for Clinical and Economic Aspects of Osteoporosis (ESCEO)/ International Osteoporosis Foundation (IOF) guidelines.[78] All costs and benefits, i.e., life years and QALYs gained, were discounted at an annual rate of 3.5% as per the NICE reference case.[1]

4.2.6 Treatment effectiveness and extrapolation

4.2.6.1 Fracture incidence

General population risk of fractures

The model inputs for the general population risk of hip, vertebral and non-hip, non-vertebral (NHNV) fractures were the same as those estimated using the method described in the IOF/ European Federation of Pharmaceutical Industry Associations (EFPIA)-endorsed study on osteoporosis in the European Union by Hernlund et al. 2013 and reported for women in various age categories from the UK in the accompanying compendium of country-specific reports by Svedbom et al. 2013.[79, 80] The incidence of hip fractures were sourced from a study by Singer et al. 1998, which was considered as the most comprehensive data on hip fracture incidence in the UK.[81] According to the company, the study by Singer reported similar findings to a more recent UK study using the Clinical Practice Research data (CPRD) link over the years 1990-2012 (i.e., van der Velde et al. 2016 which also showed that the incidence of hip fractures remained stable over the studied time period.[82] Due to unavailability of data

on the risk of clinical vertebral fractures in the UK, the incidence of vertebral fractures was estimated based on the ratio of clinical vertebral to hip fractures in a Swedish study.[83] The incidence of NHNV fractures was estimated based on a combination of the incidence of forearm fractures (distal forearm, distal radius and wrist) that was sourced from Singer et al. 1998,[81] and the ratio of "other fractures" (femur, pelvis, humerus, rib, clavicle, scapula and sternum) to hip fractures in Sweden applied to the incidence of hip fractures as estimated by Singer et al. 1998 for the UK.[80, 83] The selected inputs for incidences of hip, vertebral and NHNV fractures are displayed in Table 4.6.

Table 4.6: Incidence of fracture per 100,000 people in the UK by age

Age	Hip[81]	Vertebral[83]	NHNV[81, 84]
50–54	33	84	633
55–59	51	142	813
60–64	81	143	979
65–69	132	192	1,425
70–74	282	397	1,928
75–79	619	602	2,891
80–84	1,236	777	3,876
85+	2,255	1,061	5,958

Based on Table 18 of the CS[1]
CS = company submission; NHNV = non-hip, non-vertebral

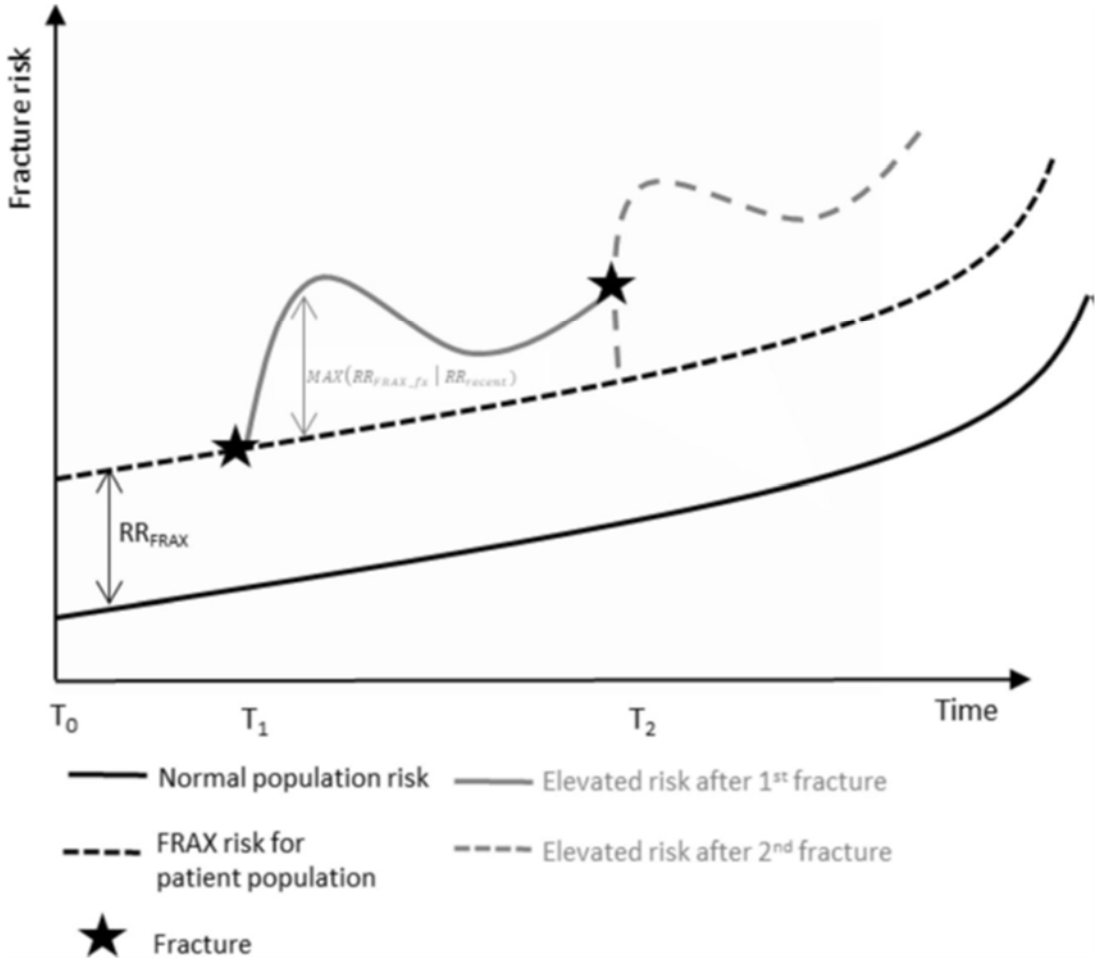
Increased fracture risk associated with osteoporosis

The model inputs for the increased fracture risk associated with osteoporosis, relative to the general population, were based on the FRAX algorithm. The FRAX tool, similar to QFracture, can be used to estimate an average 10-year risk of fracture based on clinical risk factors including age, BMI, BMD and lifestyle factors (e.g., smoking). The use of fracture risk assessment tools, such as FRAX and QFracture, in clinical practice is recommended by NICE clinical guideline (CG) 146.[11] The company preferred to use FRAX over QFracture because FRAX can be used in combination with BMD, is more widely used than QFracture, is included in the National Osteoporosis Guideline Group (NOGG) 2017 clinical guideline,[76] and can be more easily adapted to also consider the imminent fracture risk.

Imminent fracture risk

The model inputs for the imminent fracture risk, defined as the increased risk of a subsequent fracture after having sustained a first, second or third fracture, were sourced from Söreskog et al. 2020.[85] This study made use of a large dataset obtained from a retrospective real-world study in Swedish women aged 50 years and over with a fragility fracture[86], and estimated HRs for the risk of MOF in women after one, two or three fractures, relative to age- and gender-matched controls. The imminent fracture risk reaches its peak level in the first year following a fracture and then slowly declines until there is little excess risk after 5 years. When subsequent fractures occur within the timeframe of imminent risk following a prior fracture, the increases in risk may accumulate over time as "fracture cascades". An illustration of an individual patient's risk trajectory is shown in Figure 4.2 for a patient without a fracture at baseline. In contrast, the company's base-case model does assume a recent fracture at baseline.

Figure 4.2: Illustration of the fracture risk trajectory estimated using imminent risk



Based on Figure 14 in the CS,[1] which was sourced from Söreskog et al. 2020.[85]
Note: In contrast to the illustration above, the company's base-case model does assume a recent fracture at baseline.
MAX = maximum; RR_{FRAX} = relative risk estimated by FRAX for a given patient profile excluding prior fracture as a clinical risk factor; RR_{FRAX_fx} = relative risk estimated by FRAX for a given patient profile including prior fracture as a clinical risk factor; RR_{recent} = relative risk of an imminent fracture; T_0 = timepoint 0, at which the patient has no fracture history; T_1 = timepoint 1, at which the patient has sustained the first fracture; T_2 = timepoint 2, at which the patient sustained the second fracture.

Total fracture risk

For patients in the model, fracture risk was calculated as a function of the UK general population risk, the RR from FRAX for a given patient profile excluding prior fracture as a clinical risk factor, the maximum of the RR due to a recent fracture vs. no fracture (i.e. the imminent risk) or the RR from FRAX for a given patient profile including prior fracture as a clinical risk factor vs. the general population, and the risk reduction from treatment (see Section 4.2.6 of the ERG report). The formula that was used for this calculation is the following:

$$MAX(RR_{FRAX\_fx} \mid RR_{recent}) * RR_{FRAX} * General\ population\ risk * Risk\ reduction\ from\ treatment,$$

where MAX = maximum; RR_{FRAX} = relative risk estimated by FRAX for a given patient profile excluding prior fracture as a clinical risk factor; RR_{FRAX_fx} = relative risk estimated by FRAX for a given

patient profile including prior fracture as a clinical risk factor; RR_{recent} = relative risk of an imminent fracture.

Reduction of fracture risk

Efficacy estimates for romosozumab/alendronate and the comparators were applied to the above baseline fracture risks. The base-case efficacy estimates for romosozumab vs. alendronate were determined from the fracture endpoints from the ARCH study.[3] In analyses vs. other comparators, efficacy was estimated using an NMA. Treatment effects were estimated on the trial ITT population.

ARCH was considered the most relevant source of clinical evidence for modelling patients at imminent risk of fracture as it is the only study of romosozumab in women with prior fracture which includes fracture outcomes. Time-to-event analysis of fracture incidences are available from the clinical study report (CSR) for clinical fracture, non-vertebral fracture, hip fracture, and MOF. Cumulative point estimates are published for 12 and 24 months for new vertebral, clinical, non-vertebral and hip fracture types.[3]

Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH.[1] Patient-level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-dependent hazard rates were calculated for the mid-point of the model cycle. In the model, efficacy of non-vertebral fractures was applied to NHNV fractures due to lack of data on all fractures excluding both hip and vertebral. For vertebral fractures, efficacy of new vertebral fractures was calculated from the published data at 12 and 24 months.[3] Efficacy for vertebral fractures beyond month 24 is based on 24 month efficacy. The resulting non-cumulative HRs of romosozumab/alendronate vs. alendronate are displayed in Table 4.7.

Table 4.7: ARCH non-cumulative efficacy data based on parametric distributions. HR of romosozumab/alendronate vs. alendronate by time point. ITT population.

Time since treatment start (months)	HR (hip fracture)	HR (new vertebral fracture, used for vertebral fracture in the model)	HR (non-vertebral fracture, used for NHNV fracture in the model)
0–6	[redacted]	[redacted]	[redacted]
7–12	[redacted]	[redacted]	[redacted]
13–18	[redacted]	[redacted]	[redacted]
19–24	[redacted]	[redacted]	[redacted]
25–30	[redacted]	[redacted]	[redacted]
31–36	[redacted]	[redacted]	[redacted]

Based on Table 19 of the CS[1]
CS = company submission; HR = hazard ratio; ITT = intention-to-treat; NHNV = non-hip, non-vertebral

ARCH compared romosozumab/alendronate to alendronate. Therefore, ARCH provides no efficacy data vs. placebo. In the model, fracture risk reductions from treatment are applied to the general population risk. Therefore, it was necessary to transform the ARCH efficacy of romosozumab vs. alendronate to romosozumab vs. placebo. To calculate RRs for romosozumab/alendronate vs. no treatment, the HRs of romosozumab/alendronate vs. alendronate alone in Table 4.7 above were applied to RRs of alendronate vs. placebo derived from the NMA (described in Section 3.4 and below). Since HRs (Table 4.7) and RRs (from the NMA) were similar, the company assumed, given the lack of RR

data from ARCH, that these could be used interchangeably.[1] The approach of using the alendronate vs. placebo data was considered reasonable given that, according to the company, the efficacy data of alendronate vs. placebo from the CS NMA do not differ significantly from other NMAs, for example NICE's most recent NMA. A comparison of results from the NMA in the current submission compared to the NMA from NICE's most recent NMA is provided in Table 4.8 below.

Table 4.8: Comparison of results in the NMA included in this submission to the most recent NICE Assessment Group NMA

Time since treatment	Time since treatment start (months)	Hip fracture (CS ITT NMA)	Hip fracture (NICE AG NMA)[a]	Vertebral fracture (CS ITT NMA)	Vertebral fracture (NICE AG NMA)[a]	Other (NHNV) (CS ITT NMA)	Other (NHNV) (NICE AG NMA)[a]
Romosozumab/ alendronate vs. placebo	0–12	[redacted]	0.39 (0.21 to 0.72)	[redacted]	0.25 (0.15 to 0.43)	[redacted]	0.71 (0.48 to 0.85)
	13–24	[redacted]		[redacted]		[redacted]	
	25–60	[redacted]		[redacted]		[redacted]	
Alendronate vs. placebo	0–12	[redacted]	0.64 (0.45 to 0.88)	[redacted]	0.50 (0.40 to 0.64)	[redacted]	0.77 (0.64 to 0.90)
	13–24	[redacted]		[redacted]		[redacted]	
	25–60	[redacted]		[redacted]		[redacted]	
Teriparatide vs. placebo[b]	0–12	[redacted]	0.35 (0.15 to 0.73)	[redacted]	0.23 (0.16 to 0.32)	[redacted]	0.58 (0.45 to 0.76)
	13–24	[redacted]		[redacted]		[redacted]	

Based on Table 20 of the CS[1]

[a] RRs in the NICE NMA were not calculated at specific timepoints; [b] Twelve-months efficacy for hip fracture was not available for the respective comparison with teriparatide in the CS NMA; twenty-four months efficacy was therefore assumed for the first 24 months for these treatments.

AG = assessment group; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis

The NMA provided efficacy estimates up to 36 months from treatment initiation. The treatments with longer treatment durations, efficacy is extrapolated beyond 36 months until the end of the treatment duration, in line with the independent academic Assessment Group's approach in the suspended NICE multiple technology appraisal (MTA) ID901.[87] Table 4.9 presents the base-case efficacy input of romosozumab/alendronate vs. placebo, where efficacy has been calculated based on the NMA using the ITT population. A scenario analysis was also conducted using the EU-label matched NMA (described in Section 3.4 based on the results presented in Appendix D.4.4).[8] The corresponding efficacy inputs for romosozumab/alendronate vs. placebo are presented in Table 4.10.

Table 4.9: Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and NMA (ITT populations)

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	NHNV fracture
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Based on Table 21 of the CS[1] and Table 48 of the response to request for clarification.[9]
CI = confidence interval; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis.

Table 4.10: Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and scenario NMA (EU label-matched population)

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Based on Table 22 of the CS[1]

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
CI = confidence interval; CS = company submission; EU = European Union; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis				

The NMA described in Section 3.4 was used to conduct scenario analyses for romosozumab/alendronate vs. other comparators, including teriparatide, denosumab, zoledronate, risedronate and raloxifene.[1] These scenarios were based on the NMA using the ARCH and FRAME ITT population, presented in Table 4.11. For completeness, the equivalent analysis performed using the EU label-matched NMA is presented in Table 24 of the CS (data only available 12-monthly instead of 6-monthly in the base-case NMA).[1]

Table 4.11: Fracture risk ratio (95% CI), by fracture type, based on network meta-analysis (NMA, ARCH and FRAME ITT population)

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	Other fracture (NHNV)
Romosozumab/ alendronate vs. placebo	0–12[a]	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Alendronate vs. placebo	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Teriparatide vs. placebo[b]	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
Denosumab vs. placebo	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Zoledronate vs. placebo	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Risedronate vs. placebo[b]	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Raloxifene vs. placebo	0–12	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]

Based on Table 23 of the CS[1]

[a] Results from FRAME are only included at month 12; results for romosozumab/alendronate from month 13 onwards only include ARCH, as discussed in Sections 3.2.7 and 3.3.2; [b] Twelve-months efficacy for hip fracture was not available for the respective comparison with teriparatide and risedronate. Twenty-four months efficacy was therefore assumed for the first 24 months for these treatments.

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	Other fracture (NHNV)
CI = confidence interval; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis.				

Fixed effect models were used for all fracture endpoints and time periods since the deviance information criterion (DIC) was lower in the fixed effect models compared with the random effect models, as shown in Appendix D.4.3 and D.4.4.[8]

As noted above, the results from the CS NMA do not differ significantly from other NMAs (Table 4.8) according to the company.[87, 88] However, one important difference is that the CS NMA considers time-specific results, unlike previously published NMAs, which have instead assumed equal efficacy across timepoints and only considered the final efficacy time point reported in each RCT.[1] By considering fracture outcomes at specific timepoints, the CS NMA was able to consider the short and long-term comparative efficacy of each osteoporosis treatment more accurately, compared to previously published NMAs. The importance of conducting a timepoint specific NMA is illustrated throughout the NMA results presented in Section 3.4 and Appendix D.4.3 and D.4.4, where it can be seen that treatment rankings and pairwise comparisons regularly varied across different time points for the same fracture outcomes.[1, 8] This is particularly important when considering bone-building treatments, such as romosozumab, which reaches the optimal clinical performance in a relatively short duration (i.e., 12 months), providing a rapid and potent effect and demonstrating the potential to interrupt such a "fracture cascade" early in the process. The accurate consideration of short-term comparative efficacy (i.e. at Month 12) is of particularly importance for patients who have incurred a recent MOF within the past 24 months and are at imminent risk of another fragility fracture, as these patients will experience particular benefit from osteoporosis treatments with fast-acting benefits.[89]

ERG comment: During the clarification phase, the ERG requested the company to explain the extent to which fracture incidences in the UK have remained stable over time and similar to those in the Singer et al. 1998 study.[81] The company responded by referring to the study by van der Velde et al. 2016 that made use of the CPRD data from the years 1990 – 2012.[82]

In van der Velde et al. 2016, the incidence of hip fractures overall remained stable at about 35/10,000 person-years, or at about 50/10,000 person-years for women aged 75 to 79 years.[82] For women in the same age group in Singer et al. 1998 this incidence was 70.74/10,000 person-years (for women aged 70 to 74 years it was 48.5 and for women aged 80 to 84 years it was 143.72). The ERG concludes that the study by van der Velde et al. 2016 indeed confirms the stability of fracture incidence over time, but also that the incidence rates in this study are substantially lower than in the study by Singer et al. 1998.[81] As such, the validity of the incidences of hip fractures that are used in the model is uncertain.

During the clarification phase, the company explained that they had not used the estimates from Singer et al. 1998 for clinical vertebral fractures because these were deemed unrealistically low in comparison to other studies. The company indicated that that could be due to vertebral fractures being treated in other healthcare facilities than those that were included in the study. Therefore, the company estimated incidence of vertebral fractures based on the ratio of clinical vertebral to hip fractures in a Swedish study.[83] The company clarified the validity of this ratio for the UK by referring to a study by Kanis et al. 2001 that, according to the company, showed that these ratios are similar between Sweden and the UK.[90] However, the ERG notes that Kanis et al. 2001 did not include an actual comparison between the

ratios of clinical vertebral to hip fractures in Sweden and the UK. As such, the validity of the incidences of vertebral fractures that are used in the model is uncertain. In response to a request by the ERG, the company performed a scenario analysis using estimates of vertebral fracture incidence by Singer et al. 1998 that resulted in an ICER that was almost twice the value of the company's base-case ICER.

The company indicated that although the incidence of radius/ulna fractures in the UK has decreased in the year 1998 relative to preceding years, it remained stable in the years 1998 – 2012 at approximately 40/10,000 person-years in the van der Velde study. Regarding the extent of similarity for the incidences of forearm fracture between the studies by Singer et al. 1998 and van der Velde et al. 2016, the company indicated that the incidences of wrist fracture in women aged 75-79 was approximately 70/10,000 person-years in Singer et al. 1998 and approximately 50-70 per 10,000 persons-years in van der Velde et al. 2016. The ERG notes that the latter incidence refers to distal forearm fractures, which is a combination of fractures in the radius/ulna and wrist (i.e., carpal fractures). The ERG notes that in the study by van der Velde et al. 2016 the incidence of wrist fracture was stable in the years up to 1998, but has doubled in the time period 1998 to 2012 and that in the study by Singer et al. 1998 the incidence of forearm fractures (i.e., radius / ulna) was 0.68 / 10,000 person- years in women aged 75-79 years. The ERG therefore concludes that the incidence has indeed remained stable over time for radius/ulna fractures but not for wrist fractures, and that the similarity of the estimates for forearm fracture incidence is low between the two studies. The company did not comment on the similarity between the ratios of the incidence of "other fractures" relative to hip fractures in Sweden and the UK. As such, the validity of the incidences of NHNV fractures that are used in the model is uncertain.

The model uses relative risk values for the imminent fracture risk that were sourced from the study by Söreskog et al. 2020.[85] It is not clear to the ERG how the values that are used in the model correspond to those reported by Söreskog et al. 2020, which is possibly due to the use of different age categories in the paper and the model. The model also specifies values of 0 for the relative risk of a 4th fracture after a 3rd fracture, in contrast to Söreskog et al. who report non-zero values for this. No explanation was provided for this aspect; therefore, it is not clear to the ERG what the underlying rationale is for the assumed 0 values.

The 'State trace' sheet of the model provides an overview of the proportions of patients having sustained their 1st, 2nd, 3rd, and 4+ hip, vertebral or NHNV fractures. Logically, over time first a proportion of patients has their first fracture, followed by a second, et cetera. However, the proportion of patients that has their first NHNV fracture remains zero throughout the model time horizon whilst there is a non-zero proportion of patients having their second NHNV fracture from the second cycle of the model onwards. The ERG could not trace the root cause of this inconsistency.

The company has assumed that the relative risks of fracture after having had a 1st, 2nd or 3rd fracture as estimated using Swedish data are transferable to the UK. To support this assumption during the clarification phase, the company referred to the geographical proximity and similarity in quality of healthcare between Sweden and the UK and the fact that previous CE studies have made the same assumption. According to the ERG, the validity of the assumption that the relative risks of fracture are transferable between the two countries is not sufficiently justified.

In previous publications based on the same model by Söreskog et al. 2020 a limitation was noted in relation to the imminent fracture risk being possibly overestimated, because not all risk factors that are included in FRAX were available to adjust the imminent risk ratios for confounding.

To conclude, the ERG is uncertain regarding the validity of the values used for the imminent fracture risk as well as regarding their implementation in the model. In response to a request by the ERG, the

company performed a scenario analysis using only the FRAX algorithm, which includes a risk factor for prior fracture regardless of fracture recency, that resulted in an increase in the ICER, becoming more than twice the company's base-case ICER.

Treatment effect on fracture risk of romosozumab/alendronate vs. alendronate alone was calculated by reconstructing patient-level data from published Kaplan-Meier curves and then fitting parametric distributions in order to calculate time-dependent hazard rates. These (survival data) analyses are not shown in the CS. In response to clarification question B7.B, the company mentioned that the analyses were conducted internally but they are not publicly available.[9] While the methods used for the survival analyses seem appropriate, it should be emphasised that the results of such analyses were are not presented. Therefore, the ERG cannot assess whether the distributions were properly fitted and cannot explore the impact of using alternative distributions on the model results.

4.2.6.2 Persistence

Suboptimal persistence to osteoporosis medications is frequently observed in UK clinical practice, and may reduce the treatment efficacy and increase the risk of fracture compared to the reduction in fracture risk seen with optimal persistence.[1] One UK-based study (N=63,350) found that 50% of all women receiving osteoporosis treatments had discontinued treatment after six months, with 68% of all women discontinuing by the end of one year.[91]

To account for this in the model all patients were at risk of treatment discontinuation in each cycle, with discontinuation reflected in their anti-fracture treatment benefits. In the base-case, patients were assumed to be at risk of discontinuation during the first three years, after which persistence remained stable until treatment was completed, based on long-term studies indicating that discontinuation rates are highest immediately after the initiation of treatment, with discontinuation rates plateauing and remaining stable after the first year and up to five years of treatment.[92, 93] A treatment duration of five years was assumed to align with previous health economic studies and recommendations from ESCEO/IOF.[78, 94, 95]

Patients who discontinued treatment could not switch to, or restart, a treatment, due to the lack of sequential evidence in the published literature, as most RCTs have been conducted in treatment naïve patients, or required a long treatment washout period prior to enrolment.[1] For persistent patients who switch treatment within a sequence in the model, patients were assigned the probability of non-persistence corresponding to the time since the start of the treatment.

In the base-case, persistence on alendronate alone was derived from Li et al. 2012, who used the UK General Practice Research Database (GPRD) to estimate persistence on osteoporosis medications among postmenopausal women in the UK.[96] In scenarios, persistence on risedronate and raloxifene were also estimated from Li et al. 2012. Persistence on denosumab was taken from a retrospective observational study using the Swedish Prescribed Drug Register,[97] while persistence on teriparatide and zoledronate were taken from a Swedish osteoporosis database.[75]

Persistence on romosozumab in clinical practice is unknown. As a starting point the company considered the persistence on teriparatide. A Swedish osteoporosis database reported that teriparatide had a 6-month and 12-month persistence of approximately 74% and 61%, respectively.[75] The company argue that as romosozumab will be administered much less frequently compared to teriparatide (QM vs QD), and UCB will provide a PSP in the UK, it is reasonable to assume that persistence on romosozumab will be higher than on teriparatide.[1] However, the size of this improvement is unknown. Based on the three pivotal romosozumab clinical trials,[3, 17, 18] the company assumed that 90% of patients

will be persistent to treatment throughout the 12-month romosozumab treatment period. In ARCH, [redacted]% of patients receiving romosozumab completed the first 12-month treatment period.

For the treatment sequence of romosozumab followed by alendronate used in this submission, it was assumed that the persistence rates for alendronate would be 85% of the persistence for denosumab. This was based on the assumption that patients who initially demonstrated high persistence on romosozumab would be expected to demonstrate high persistence on follow-on treatments, and therefore the persistence on alendronate after romosozumab would be notably higher than the persistence on alendronate alone reported by Li et al. 2012.[96] The company report that this assumption is supported by a study of persistence to treatment in chronic diseases, which found that patients who have already persisted on treatment for a year have a 50% reduced discontinuation rate compared to patients just starting treatment.[98] Additionally the company note that the patient population in Li et al. 2012 is less severe that the target population for romosozumab, as they were not required to have experienced a previous fracture, while patients eligible for treatment with romosozumab/alendronate will have experienced a recent MOF within 24 months.[96] The company would expect that these more severe patients would exhibit improved persistence and that USB's PSP will include support with the transition to follow-on treatment, which is likely to further increase persistence on alendronate after romosozumab compared to alendronate alone.[1] A summary of persistence assumptions for all treatments can be found in Table 4.12.

Table 4.12: Proportion of patients on osteoporosis treatment over time in the economic model

Month since treatment initiation	Romosozumab	Alendronate after romosozumab[a]	Alendronate alone[96]	Teriparatide[b75]	Zoledronate[b75]	Denosumab[b97]	Risedronate[b96]	Raloxifene[b96]
6	90%	85%	49%	74%	100%	100%	50%	45%
12	90%	71%	38%	61%	100%	83%	38%	33%
18	0%	59%	34%	3%	51%	69%	33%	30%
24	0%	53%	30%	3%	42%	62%	28%	26%
30	0%	47%	27%	0%	34%	56%	24%	23%
36	0%	43%	24%	0%	28%	50%	21%	21%
42	0%	38%	22%	0%	23%	45%	18%	19%
48	0%	34%	20%	0%	18%	40%	16%	17%
54	0%	31%	19%	0%	15%	36%	14%	16%
60	0%	28%	17%	0%	12%	33%	12%	14%

Based on Table 25 of the CS[1]
[a] The persistence on alendronate after romosozumab was assumed to be 85% of the persistence on denosumab;
[b] Treatment included in scenario analyses only.
CS = company submission

ERG comment: The company's approach to model persistence is inconsistent between intervention (romosozumab) and comparators and is likely to be biased in favour of romosozumab. The guidelines for economic evaluations in osteoporosis endorsed by the ESCEO/IOF recommend using real-world data on medication adherence.[99] However, this approach was only used for the comparators.

The company assumed that persistence with romosozumab is 90%, which was based on persistence with romosozumab as observed in the ARCH trial. However, in response to clarification question B9 the company indicated that "*persistence data from retrospective observational studies are more appropriate than persistence data from clinical trials. Persistence in clinical trials is significantly higher than in clinical practice most likely because patients know they are being observed and have consented to participate in the study*" and that "*persistence of romosozumab is assumed to be the same as in the ARCH trial, despite clinical trials show higher persistence than what is seen in clinical practice. This was necessary given that there is no real-world evidence currently available for romosozumab as it has only been recently launched*".[9] The ERG agrees that real-world persistence with romosozumab, outside the context of a clinical trial, will be lower than in ARCH and therefore prefers to use a lower value for their base-case analysis. In line with the assumption made by Söreskog et al. 2021 in their CE analysis for romosozumab in Sweden,[73] the ERG assumes a value of 80% for persistence with romosozumab. The ERG considers this a plausible value since it is lower than persistence with romosozumab in ARCH and higher than the real-world persistence with teriparatide that the company sourced from the Swedish osteoporosis database. The latter is supported by the notion that romosozumab will be administered less frequently than teriparatide and that it is likely that persistence with romosozumab is higher relative to treatments with higher frequencies of administration.

For persistence with alendronate, the company assumed lower values for persistence with alendronate alone than for persistence with alendronate after romosozumab. Specifically, the company assumed that persistence with alendronate after romosozumab is 85% of persistence with denosumab as sourced from a Swedish study by Karlsson et al. 2015.[97] The ERG considers this an arbitrary choice. The company sourced persistence with alendronate alone from Li et al. 2012,[96] which was a study on persistence with osteoporosis therapies based on UK CPRD data. The company justified the use of different sources by referring to a difference in the severity of osteoporosis between patients treated with either alendronate after romosozumab or alendronate alone. Since alendronate, as a standalone treatment, was positioned as the most relevant comparator to romosozumab in the indicated population for the company's base-case analysis, the ERG considers it inappropriate to assume a difference in severity of osteoporosis for the population that is considered eligible for both treatment options. Therefore, the ERG prefers to inform persistence with alendronate, regardless of whether it is given as a standalone treatment or after romosozumab, using the same study. Furthermore, the ERG was unable to verify the persistence values shown in Table 4.12 that the company indicated were sourced from Li et al. 2012.[96] Importantly, the data in the study by Li et al. 2012 range from 1995 to 2008 and indicate that persistence estimates have not been stable over that period of time. The ERG identified a more recent study by Morley et al. 2020 on persistence with osteoporosis therapies that also made use of UK CPRD data.[100] The ERG preferred to use this more recent source of persistence estimates for their base-case.

In addition to persistence with alendronate after romosozumab and alendronate alone, the ERG also used the study by Morley et al. 2020 to inform persistence with denosumab, risedronate and raloxifene using data from the subgroup of naïve patients.[100] Whilst Morley et al. 2020 also provide estimates for persistence with teriparatide and zoledronate, the ERG did not use these estimates because they were based on very small (n<20) sample sizes. Instead, the ERG preferred to use the same estimates as the company for persistence with these comparators. However, the ERG did not have access to the Swedish osteoporosis database that informed these estimates nor any details regarding the methods that were used. As such, the validity of these estimates remains uncertain. The ERG preferred estimates of persistence are presented in Table 4.13.

Table 4.13: ERG preferred estimates of persistence with osteoporosis therapies

Month since treatment initiation	Romosozumab[73]	Alendronate after romosozumab[100]	Alendronate alone[100]	Teriparatide[a,b]	Zoledronate[a,b]	Denosumab[a100]	Risedronate[a100]	Raloxifene[a100]
6	80%	31%	62%	74%	100%	64%	62%	53%
12	80%	19%	51%	61%	100%	55%	51%	42%
18	0%	14%	44%	50%	51%	48%	44%	37%
24	0%	11%	38%	41%	42%	36%	38%	33%
30	0%	9%	34%	0%	34%	32%	34%	29%
36	0%	8%	29%	0%	28%	28%	29%	25%
42	0%	7%	26%	0%	23%	25%	26%	24%
48	0%	6%	24%	0%	18%	22%	24%	24%
54	0%	5%	21%	0%	15%	19%	21%	23%
60	0%	4%	18%	0%	12%	16%[c]	18%	22%

[a] Treatment included in scenario analyses only; [b] Same values as company base-case; [c] In absence of value for naïve patients, the value from 'All patients' was used.

The company indicated in the CS that differences in persistence exist between patients that previously persisted on osteoporosis treatment (i.e., non-naïve patients) and patients that just started with osteoporosis treatment (i.e., naïve patients). For example, Morley et al. 2020 found that persistence with oral bisphosphonates was higher in naïve patients than in non-naïve patients.[100] This contrasts with findings from an earlier study that found the opposite.[101] Also, the company considered that the PSP is likely to increase persistence on alendronate after romosozumab compared to alendronate alone. However, this assumption is not based on any evidence. To address the uncertainty surrounding this aspect and the extent to which patients can still be considered as naïve once they have persisted with a six-month treatment course, the ERG assessed the impact on the CE results when assuming the same persistence for naïve and non-naïve patients. For this the ERG performed a scenario analysis in which persistence was based on the pooled data from all patients (i.e., both naïve and non-naïve patients) in Morley et al. 2020, for both alendronate alone and for alendronate after romosozumab.[100]

4.2.6.3 Dynamic residual effects

The company assume that the time a patient remains on osteoporosis treatment is directly related to the duration of efficacy that can be expected. They argue that there is consensus that anti-fracture efficacy persists for a period of time (offset time) after treatment is discontinued in patients with osteoporosis.[102] Two alternatives for modelling residual effects are presented in the CS and in Figure 4.3 below:[1]

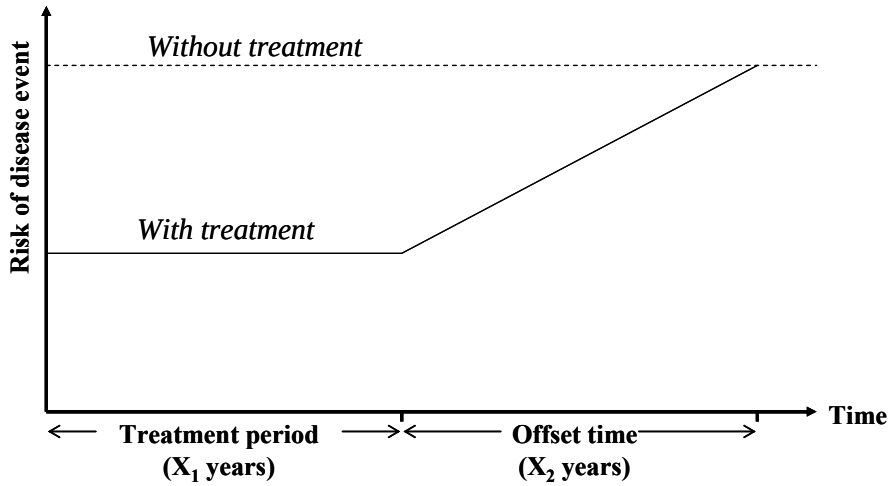
- Dynamic: Offset time is assumed to be as long as time on treatment and is, therefore, shorter for patients who drop out earlier. Partially persistent patients are distributed over a range of treatment durations and corresponding offset times depending on if and when they stopped treatment.

- Fixed: All patients have the same specified offset time irrespective of treatment drop out, so a patient who discontinues after 1 year will nonetheless have 2-years offset time if the prespecified offset time was 2 years.

During the offset time the fracture risk reduction is assumed to decline linearly to zero.[1] The efficacy of the last treatment given to the patient in the sequence was used for the offset time. Thus, if a patient was treated with romosozumab for 12 months and alendronate for the following 36 months, the offset time equalled 48 months and efficacy used for offset was based on the efficacy of alendronate for patients who had previously received romosozumab. This was validated by leading UK experts at an advisory board. This approach is recommended by the ESCEO and IOF guidelines, and has been used in other published health economic studies and romosozumab HTA submissions to the Scottish Medicines Consortium (SMC) and TLV (Tandvårds- och läkemedelsförmånsverket, The Swedish Dental and Pharmaceutical Benefits Agency).[78, 103, 104]

The company report that evidence supports the assumption that alendronate, zoledronate and teriparatide have offset times similar to the treatment length and there is no robust evidence to support differential offsets for other treatments, providing evidence for the dynamic model approach.[105-109] For denosumab, efficacy was limited to 6 months after discontinuation.[110, 111] Chronic treatment with denosumab is necessary when used as the subsequent treatment after romosozumab for this combination to provide optimal benefits to patients; or alternatively a further treatment switch to a bisphosphonate after the denosumab treatment period would be required. In the model, a one-year fixed offset time was applied to denosumab.[1] This was described by the company as a conservative approach. A summary of the treatment sequences and associated length of effects is presented in Table 4.14 (a complete description of the scenarios is given in Section 5.2.3).

Figure 4.3: Modelling the residual effects of osteoporosis treatments



Based on Figure 15 of the CS.[1]
CS = company submission; X_1 = treatment period; X_2 = offset time

ERG comment: The company assumptions regarding dynamic residual treatment effects are broadly in line with the recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019.[78] Therefore, the ERG considers the company's approach appropriate. Scenarios with fixed offset time can be deemed as exploratory.

As described in Key issue 2, a scenario analysis where treatment waning starts at four years followed by a dynamic offset (linear waning) of the treatment effect was explored by the ERG in Section 6.1.2.

In this scenario it was assumed 4 years of full effect, a waning in effect for one more year (the waning assumption was to consider an effect between sequential alendronate and alendronate alone as assumed by the company) followed by a dynamic offset 5 years. Note, however, that the other scenario mentioned in Key issue 2, one with shorter duration of the dynamic offset of the treatment effect was not possible to run. In the model implementation, offset time is either dynamic and equal to the time on treatment, or fixed to 1 year. The rationale for the second scenario was that, if treatment effect waning is possible, the duration of the residual treatment effect might be less than the time on treatment. Thus, for the combination romosozumab/alendronate, the ERG wanted to explore a scenario where the offset time was three years instead of the five assumed in the model. The ERG was unable to run this scenario, which is expected to increase the ICER.

Finally, the ERG would like to note that residual effects for zoledronate could be longer than those assumed by the company.[112] However, the ERG was unable to change the model to incorporate this assumption. Cost effectiveness results including zoledronate as comparator might be underestimating the ICER.

Table 4.14: Summary of treatment sequences and treatment effect duration applied for the base-case and company scenario analyses

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Base-case scenario											
Intervention: ROMO/ALN	ROMO	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Comparator: ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 1											
ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 2											
TRP (24 months)	TRP	TRP	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect
Scenario 3											
TRP (18 months)	TRP	TRP (1/2) NONE (1/2)	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Scenario 4											
TRP (biosimilar)/ALN	TRP	TRP (1/2) ALN (1/2)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 5											
TRP/ALN	TRP	TRP (1/2) ALN (1/2)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 6											
RAL	RAL	RAL	RAL	RAL	RAL	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 7											
DEN	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect
Scenario 8											
RIS	RIS	RIS	RIS	RIS	RIS	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 9											
ZOL	ZOL	ZOL	ZOL	ZOL	ZOL	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Scenario 10											
ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 11											
DEN	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect

Based on Table 43 of the response to request for clarification (question B1).[9]
*Treatment effect on fracture risk reduction.
ALN = alendronate, DEN = denosumab, RAL = raloxifene, RIS = risedronate, ROMO = romosozumab, TRP = teriparatide, Tx. = treatment, ZOL = zoledronate

4.2.6.4 Mortality

Mortality is captured in the model in three ways: age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor.[1] Age- and gender-specific mortality rates for the general population (all-cause mortality) in the UK were based on the years 2012–2014.[113] At the start of the model mortality risk is determined by UK general population all-cause mortality. When a patient sustains a fracture, the relative risk of death compared with the non-fractured population is applied to the normal population risk, and the relative risk was down-adjusted to 30% to adjust for higher frailty (i.e., increased risk of death due to other reasons than the fracture itself) in the fractured population.[1, 94, 114]

ERG comment: It is unclear why the company used UK Life Tables from 2012 to 2014.[113] In the ERG base-case, the most recent version (2017 to 2019) was used.[115]

Mortality related to hip and clinical vertebral fractures

For hip fractures, age-dependent relative risks of death were sourced from Jönssen et al. 2011,[103] a study on the CE of denosumab in Sweden. The estimated mortality during the first and subsequent years after hip fracture from a sample of 36,551 Swedish women with a main diagnosis of femur fracture between 1997 and 2001 were used to calculate standardised mortality ratios (SMRs) relative to the mortality of the Swedish age- and gender-matched general population in 2000. It was assumed that the SMRs based on Swedish data were generalisable to the UK. For vertebral fractures, the age-dependent relative risks of death were also sourced from Jönssen et al. 2011.[103] In that study, mortality was based on data from a Swedish sample that included 994 patients who sustained a clinical vertebral fracture in 1993 to 1994.[116] The age- and sex-dependent mortality was used to calculate SMRs in the same way as for hip fractures, but relative to the mortality of the Swedish general population in 1994. The relative risks of mortality compared to the normal population are presented in Table 4.15 below.

Table 4.15: Relative risk of mortality for hip and clinical vertebral fractures compared to the general population

Age	Hip fracture Year 1[103]	Clinical vertebral fracture Year 1[116]	Hip fracture Year 2+[103]	Clinical vertebral fracture Year 2+[116]
50 years	9.79	12.07	3.62	7.94
55 years	8.64	10.15	3.34	6.67
60 years	7.69	9.04	3.11	5.94
65 years	6.39	7.43	2.70	4.88
70 years	5.54	5.98	2.44	3.93
75 years	4.16	4.39	1.91	2.88
80 years	2.92	2.75	1.39	1.81
85 years	2.15	1.98	1.06	1.30
90 years	1.63	1.36	1.00	1.00

Based on Table 26 of the CS[1]
CS = company submission

Mortality relating to NHNV fractures

For NHNV fractures, the relative risk of death was calculated as a weighted average of the estimates of relative risks reported by Barret et al. 2003 using the proportions of different fracture types reported by Kanis et al. 2001.[90, 117] The company assumed that the relative risks of death after rib (30% of the included fractures) and clavicle/scapula/sternum (13% of the included fractures), which were not reported by Barret et al. 2003 were equal to one (i.e., no excess mortality). The same relative risk was used for all ages, which the company justified by referring to the variation in fracture distribution across age groups which was deemed to be small. The company notes that since the relative risk of death for NHNV fractures is known to increase with age,[116, 118, 119] the use of the same estimate for all age groups could lead to underestimation in younger and overestimation in older patients. The estimated mortality after NHNV fracture is shown in Table 4.16. It was assumed that women sustaining a fracture at NHNV sites were at increased risk of death only within the first year of fracture.

Table 4.16: Mortality during the first year following NHNV fractures

Fracture type	Fractures	Proportion	Relative risk of death
Rib	340	30%	1.0
Pelvis	47	4%	1.7
Proximal humerus	352	31%	1.4
Humeral shaft	117	10%	1.2
Clavicle, scapula, sternum[a]	145	13%	1.0
Other femoral	52	5%	1.8
Tibia, fibula	98	9%	1.1
All	1,151	100%	1.23

Based on Table 27 of the CS[1]
[a] No excess mortality reported, relative risk assumed to be equal to 1.0.
CS = company submission; NHNV = non-hip, non-vertebral

Comorbidity adjustment excess mortality

It has been reported that patients with osteoporosis have a higher degree of frailty compared to the general population and that excess mortality after a fragility fracture is not entirely attributable to the fracture event. A common assumption is that 30% of excess mortality is directly caused by the fragility fracture.[94, 114] Therefore, it was assumed that 30% of excess mortality after hip, clinical vertebral or NHNV fracture was associated with the fracture event.

The model also assumed that a patient would incur the highest risk of excess mortality, depending on previous fracture history. For example, if a patient sustained a hip fracture in cycle three and an NHNV fracture in cycle five, the excess mortality risk that was highest was incorporated (in this instance the second-year hip fracture excess mortality). The increased mortality was assumed to persist for 8 years, in line with the follow-up period in previous studies.[119, 120]

ERG comment: For the calculation of the relative risk of death for NHNV fractures, the company used the incidence of fractures for the age group of 65 to 69 years from Kanis et al. 2001.[90] The ERG notes that the incidence an older age group (e.g., 70 to 74 years or 75 to 79 years) would have made for a better match with the modelled population, but this is unlikely to have a substantial impact on the CE results.

The ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] suggest that only the excess mortality of hip and vertebral fractures should be included, as there is not yet enough evidence regarding NHNV fractures. However, there was a lack of consensus on this inclusion of excess mortality due to vertebral fractures amongst the 23 clinical and economic experts that were asked to review and validate the recommendations. In light of this, the ERG prefers to include excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNVfractures were also explored by the ERG in Section 6.1.2.

Modelling mortality with the FRAX algorithm

Some of the clinical risk factors that are inputted into the FRAX algorithm are known to contribute to mortality. Based on this, one of the outputs of the FRAX algorithm is the relative risk of pre-fracture mortality for the defined patient population.[1] This relative risk was used to adjust the baseline mortality of patients in the model, as well as mortality after fracture. However, this assumed that the pre-fracture relative risk of mortality obtained from FRAX did not change once a patient had experienced a fracture. This assumption was made as the relationship between clinical risk factors and mortality post-fragility fracture has not yet been investigated.[1]

Using mortality relative risks from the FRAX algorithm resulted in higher risk populations having a higher overall mortality (compared to lower risk populations), and thus benefiting less from avoiding fractures, compared to if the mortality adjustment was not included.[1]

The FRAX algorithm does not take into account other risk factors (not inputted into the FRAX algorithm) that may differentiate the mortality of osteoporosis patients compared to the general population. Consequently, the assumption that only a proportion of the excess mortality after fracture is related to the fracture event is made, as described above. The model uses the highest mortality in situations where both post-fracture mortality and FRAX-derived mortality need to be accounted for.

4.2.7 Adverse events

The company note several AEs that can be associated with osteoporosis regimens include upper gastrointestinal (GI) symptoms, osteonecrosis of the jaw (ONJ), hypocalcaemia, bone pain, atypical femoral fractures (AFFs), influenza-like symptoms, conjunctivitis, atrial fibrillation and stroke.[1] However, they report that due to lack of evidence, the model only includes gastrointestinal adverse events (GIAEs) that are associated with oral bisphosphonates, and excludes other AEs associated with osteoporosis, in line with other economic models and previous NICE appraisals of anti-osteoporotic treatments.[121, 122] The CS confirmed that no adjudicated events of ONJ or AFF were reported in the 12-month double-blind ARCH treatment phase.[3] During the open-label alendronate treatment phase, only one ONJ event occurred in each arm (<0.1% each in the alendronate/romosozumab and alendronate/alendronate arms) and six AFF events (two events (<0.1%) and four events (0.2%) respectively) were observed.[1]

An imbalance in serious adjudicated CV AEs was observed in the ARCH trial.[1] Romosozumab is therefore contraindicated for patients with previous myocardial infarction (MI) or stroke.[4] Given this contraindication, which was not an exclusion criterion in the ARCH trial, the company considered it reasonable to exclude CV AEs from the economic analysis. They stated that this approach aligned with the independent academic Assessment Group's approach in the suspended NICE MTA ID901.[87]

ERG comment: It was unclear whether all CV events in the ARCH trial occurred in individuals with a history of MI or stoke. If not, then the exclusion of those events which occurred in people who would not be contraindicated would be inappropriate. At clarification the ERG requested that the company

included CV AEs in the model according to the incidence in the ARCH trial.[123] In response, the company included a scenario utilising the relative risk of a CV-event based on the ARCH study, including only patients who do not have the contraindication of prior MI or stroke.[9] The post-hoc analysis of ARCH showed that patients randomised to romosozumab who did not have the contraindication (MI or stroke) at baseline, had a relative risk of major adverse CV events of [redacted] during the first [redacted] years after randomisation, compared with alendronate (subject incidence [redacted]% in romosozumab arm vs. [redacted]% in alendronate arm).[9] Costs and disutilities related to CV events are described in the relevant HRQoL and cost sections.

4.2.8 Health-related quality of life

4.2.8.1 Health state utility values

HRQoL was assessed in the ARCH trial at pre-determined time points, irrespective of fracture occurrence. The company considered it inappropriate to use this trial QoL data as it did not provide robust sensitive utility values for fracture health states.[1] The collected QoL data were also treatment specific, which the company expected would underestimate the potential QoL gain associated with treatment.

Therefore, the company preferred to use utility multipliers for fractures from the International Costs and Utilities Related to Osteoporotic Fractures Study (ICUROS) combined with UK general population values from Szende et al. 2014.[73, 74, 124] The ICUROS study was designed to assess the QoL impact of fractures on osteoporosis patients over time for use in CE modelling. It is the largest prospective study on osteoporosis quality by including over 7,000 patients in 12 countries, including 357 fractures experienced by patients in the UK.[73, 74] The ICUROS measured QoL using the EQ-5D as soon as possible after fracture occurrence regardless of treatment, and then at 4, 12 and 18 months after fracture, allowing the estimation of short- and long-term impact of osteoporotic fracture in real-world patients. ICUROS utilities were used by the independent Assessment Group in technology appraisal (TA) 464 and have also been used in economic evaluations of romosozumab for the TLV in Sweden and the SMC in Scotland.[11, 104, 125] The ESCEO/IOF also recommend using national ICUROS data if available or otherwise the international version. The utility multipliers for the first year after fracture and the second and following years are displayed in Table 4.17.

Table 4.17: Utility multipliers

Health state	Multiplier	Reference
First year after fracture		
Hip fracture	[redacted]	ICUROS
Vertebral fracture	[redacted]	ICUROS
Other NHNV fractures	[redacted]	ICUROS
Second and following years after fracture		
Hip fracture	[redacted]	ICUROS
Vertebral fracture	[redacted]	ICUROS
Other NHNV fractures	[redacted]	ICUROS

Based on Table 28 of the CS[1]
CS = company submission; ICUROS = International Costs and Utilities Related to Osteoporotic Fractures Study; NHNV = non-hip, non-vertebral

These multipliers were applied to the UK general population utility values estimated by Szende et al. 2014 shown in Table 4.18.[124] Disutilities for multiple fractures were applied in a multiplicative approach.

Table 4.18: UK General population utility values

Age	General population utility
50 years	0.849
55 years	0.804
60 years	0.804
65 years	0.785
70 years	0.785
75 years	0.734
80 years	0.734
Source: Table 29 of the CS[1] CS = company submission	

ERG comment: The ERG agrees with the approach of using fracture event utility multipliers from a large study rather than the ARCH data, which was collected at set times rather than on occurrence of fracture events. The ICUROS study included patient data from EuroQoL-5 Dimensions-3 Levels (EQ-5D-3L), time trade-off (TTO) and EuroQoL-Visual analogue scale (EQ-VAS). In the clarification response, the company clarified that multipliers were based on EQ-5D-3L data only, not the TTO or EQ-VAS data.[9] This aligns with the measurement aspect of the NICE reference case. In their clarification response the company also clarified that the utility multipliers obtained from the ICUROS study were based on data from all countries included in the study as UK specific multipliers are not currently available.[9] However, it would appear, given the similarity of the current multipliers with those used in ID901 (shown below) that the UK value set, which was used in ID901, was also used to estimate utility multipliers in this case.[87] Therefore, while utilities may be slightly affected by different reporting of health in different countries (for example due to different quality of treatment or interpretation of response options), utilities are not affected by different preferences across countries as the UK value set was used for all countries. This increases the likelihood that values are representative of UK utilities.

The multipliers included in this submission differ somewhat from those used in TA464 and ID901, as shown in Table 4.19.[11, 87] ID901 multipliers are fairly similar to those presented in this submission. However, the multipliers presented in TA464 suggest that hip and NHNV fractures have less impact on HRQoL compared to the current submission, while vertebral fractures have more impact. The company stated at clarification that the difference between the current submission and ID901 in NHNV fractures was due to the fact that UCB included more fracture types than ID901.[9] Detailed data from ICUROS on utilities for additional fracture types were found in the appendix of a study by Kanis et al. 2018.[126] Other differences with TA464 were considered to be due to the larger sample size available in the analysis by the company, which included around 3,000 fracture patients rather than just over 1,000 in the prior appraisal. These alternative sets of multipliers will be considered in a scenario to explore the sensitivity of results to multipliers used.

Table 4.19: Utility multipliers across submissions

Health state	ID3936	ID901	TA464
First year after fracture			
Hip fracture	[redacted]	0.55	0.69

Health state	ID3936	ID901	TA464
Vertebral fracture	[redacted]	0.68	0.57
Other NHNV fractures	[redacted]	0.805*	0.87**
Second and following years after fracture			
Hip fracture	[redacted]	0.86	0.85
Vertebral fracture	[redacted]	0.85	0.66
Other NHNV fractures	[redacted]	0.995*	0.99**

Based on CS[1], NICE TA464,[11], and AG report[87]

* ID901 provided multipliers for proximal humerus and wrist separately. The multipliers in the table above have been estimated as the mean of the proximal humerus and wrist values presented (year 1, 0.78+0.83/2 = 0.805; and year 2, 1.00+0.99/2 = 0.995); ** TA464 provided multipliers for shoulder and wrist separately. The multipliers in the table above have been estimated as the mean of the shoulder and wrist values presented. (year 1, 0.86+0.88/2 = 0.87; and year 2, 1.00+0.98/2=0.99)

AG = assessment group; CS = company submission; NHNV = non-hip, non-vertebral; NICE = National Institute for Health and Care Excellence; TA = technology appraisal

The multiplicative approach for accounting for the impact of multiple chronic or acute fractures has been used in previous appraisals.[11, 87] The way in which chronic multipliers were combined differs somewhat across appraisals. In TA464, if more than one fracture occurred then the chronic multipliers for each fracture was applied, but no more than one acute fracture was applied at any one time.[11] In their clarification response the company confirmed that they assumed that a maximum of two acute multipliers could be applied at once.[9] It is unclear which approach is more appropriate in this case, but the ERG could not test the impact of this assumption as changing the VBA code was not possible.

The ERG felt it was important to understand how long we would expect these chronic multipliers to continue for and whether it is realistic that the relative impact of a fracture on HRQoL at 2 years will be the same as the impact at 10 years. The company reported evidence of long-term impact of fractures from several studies in response to clarification question B17D.[9] This included studies by Adachi et al. 2011, Blomfeldt et al. 2005 and Ekström et al. 2009.[127-129] These studies found that EQ-5D utilities remained lower than pre-fracture utilities after 3-, 5- and 2-years post-fracture, respectively.[127-129] Although the ERG could only see evidence up to 4 years in the Blomfeldt publication, it did show a continuing steady decline in utility between months 4, 12, 24, and 48 post-displaced femoral neck fracture, which could be likely to continue.[128] Ekström shows a steady-state lower post-fracture utility at months 4, 12 and 24 post- subtrochanteric fracture.[129] These studies suggest that a long-term effect of fracture on HRQoL could be appropriate. The same lifetime chronic multiplier assumption was made in TA464 and ID901, so could be considered an accepted approach. The ERG could not test the impact of this assumption as they could not change the VBA code in the model and the company declined to add an option for a reduced duration of chronic multipliers in the model.

4.2.8.2 Disutility values

Utility decrements were included for patients experiencing GIAEs whilst on oral bisphosphonate treatment. A fixed QALY decrement of 0.0075 was applied at the start of the treatment without adjustment for baseline health utility for 3% of patients when starting treatment with an oral bisphosphonate, in line with the assumptions included in Davis et al. 2015 as part of NICE TA464.[95]

ERG comment: It is unclear how this disutility was calculated in TA464 but given the size of the disutility and the percentage of patients it is applied to it is unlikely to have a large impact on results.

At clarification the company provided the option to include CV AEs in the model. A multiplier for QoL after a CV event was estimated based on a Swedish study by Lindgren et al. 2007,[130] which estimated a QoL loss of 0.075 (multiplier 0.910) during the first year after CV event. For the second and following years, the multiplier was assumed to be 0.95 due to lack of data.[9]

4.2.9 Resources and costs

The following cost categories were included in the analysis: drug acquisition costs, drug administration costs, disease management costs, costs associated with fractures (i.e., hip fractures, vertebral fractures, and NHNV fractures), long-term care costs after a hip fracture, and costs for the treatment of GIAEs.

4.2.9.1 Drug acquisition costs

The drug acquisition costs for romosozumab are £427.75 per set of two pre-filled disposable 1.17 ml injections of 90 mg/ml at list price or [redacted] including the Patient Access Scheme (PAS) discount, resulting in an annual cost of £5,133 at list price, or [redacted] including the PAS discount. The drug acquisition cost for alendronate at list price is £0.96 per pack with four tablets of 70 mg, or £13 annually. The cost of the comparators used in the scenario analyses are provided in Table 4.20 below.

Table 4.20: Drug acquisition costs

Drug	Annual drug cost	Pack size and cost	Method of administration	Dosing interval	Source
Treatments used in base-case analysis					
Romosozumab[a]	List: £5,133 PAS: £[redacted]	Injection, 90 mg/ml, consisting of two pre-filled disposable injections List: £427.75 PAS: £[redacted]	SC	QM	BNF 2021,[131] PAS
Alendronate	£13	70mg 4-tablet pack (£0.96)	Oral	QW	BNF 2021[131]
Treatments used in scenario analyses					
Teriparatide[b] (Forsteo)	£3,547	Injection, 250 micrograms/ ml, net price 2.4 ml prefilled pen=£271.88	SC	1 day	NHS indicative price 2021
Teriparatide[b] (Movymia)	£3,065	Injection, 250 micrograms/ ml, net price 2.4 ml prefilled pen (£235)	SC	QD	NHS indicative price 2021
Denosumab	£371	One pre-filled disposable injection (£180)	SC	Q6M[c]	BNF 2021[131]
Risedronate	£68	35mg 4-tablet pack (£18.88)	Oral	QW	BNF 2021[131]
Zoledronate	£85	Generic zoledronate 5[c] mg/ 100ml infusion bag	IV	Yearly	BNF 2021[131]
Raloxifene	£50	28-tablet pack (£3.81)	Oral	QD	BNF 2021[131]

Based on Table 31 in the CS.[1]

[a] Romosozumab is a 12 month course of treatment; [b] Treatment with teriparatide is limited to 24 months during a lifetime.[132]; [c] The ERG corrected the information from the CS, as explained in the ERG comment in Section 4.2.4 of the ERG report.

BNF = British National Formulary; CS = company submission; IV = intravenous; NHS = National Health Service; PAS = Patient Access Scheme; QD = once daily; QM = once monthly; Q6M = once every 6 months; QW = once weekly; SC = subcutaneous

4.2.9.2 Drug administration costs

No drug administration costs were included for romosozumab, which the company justified by referring to their plans to set up a Patient Support Programme (PSP) that includes homecare service, an adherence support program, and training of injection techniques. Administration costs are not included for alendronate since it is administered orally.

Drug administration costs were included in the model only for patients receiving denosumab or zoledronate. For patients receiving denosumab these consist of two nurse visits per year, which were valued at £9.50 assuming a 15 minute visit and using a unit cost of £38 per hour as provided by the Personal Social Services Research Unit (PSSRU) 2020.[133] For patients receiving zoledronate the administration cost was valued at £160 assuming the same cost as for delivery of chemotherapy and using the NHS National Tariff Workbook 2020/2021 (HRG code SB12Z; Deliver Simple Parental Chemotherapy at First Attendance).[134]

ERG comment: During the clarification phase, the ERG requested the inclusion of administration costs for romosozumab (i.e., representing a situation where the PSP is not in place) and all relevant comparators. The company responded by providing the results of scenario analyses that included the following administration costs in addition to those included in the original analyses: 12 nurse visits per year for romosozumab and 365 nurse visits per year for teriparatide. Nurse visits were valued at £9.50 (i.e., the same as above). For their base-case analysis, the ERG assumed a situation where the PSP has not (yet) been implemented and includes the costs for administration (i.e., 12 nurse visits) of romosozumab. The ERG performed a scenario analysis where it is assumed that the PSP is in place, and in which the costs of administration are applied in isolation as well as in combination with the assumption that persistence with romosozumab is 90%. The latter scenario was included since it is likely that the PSP leads to improvements in persistence with romosozumab.

4.2.9.3 Disease management costs

Disease management costs that were included in the model consist of BMD measurements and physician (GP) visits. BMD measurements were modelled at a frequency of once per two years and were valued at £40 using the NHS National Tariff Workbook 2020/2021 (RD50Z, DXA scan).[134] Physician visits for the monitoring of osteoporosis therapies were modelled at a frequency of once per year and were valued at £39 using the unit cost for a 9.22 minutes consultation as provided by the PSSRU 2020.[133]

ERG comment: The inclusion of costs for BMD measurements and physician visits was in line with Borgström et al. 2006 and Jönssen et al. 2011.[103, 135] However, other economic evaluations have included the costs of physician visits at a frequency of twice per year instead of only once, as indicated in the ESCEO/IOF recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019 and as used in Hiligsmann et al. 2020. The ERG preferred base-case analysis therefore assumed a frequency of twice per year for physician visits.

4.2.9.4 Fracture costs

The costs of hip, vertebral, and NHNV fractures during the first year after a fracture were sourced from a study by Gutiérrez et al.,[136, 137] and updated to 2020 using the consumer price indices (CPIs) as provided by the Office for National Statistics (ONS).[138] This resulted in cost estimates of £13,203, £2,897, and £2,131 for the first year after a hip, vertebral, or NHNV fracture, respectively. The costs of fractures in subsequent years were sourced from Davis et al. 2016,[95] and updated to 2020 using the CPIs as provided by the ONS.[138] These were only applied to hip and vertebral fractures at £115 and £361,

respectively. The costs of long-term care were included as recommended by the ESCEO/IOF recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019 and in line with TA464.[11, 78] In line with TA464, the probabilities of discharge to institutional care by age group were sourced from Najayan et al. 2014.[11, 117] The cost of long-term care in a nursing home was sourced from Hernlund et al. 2013,[80] and updated to 2020 using the CPIs as provided by the ONS,[138] which resulted in a daily cost of £112.

ERG comment: The first-year costs of hip, vertebral and NHNV fractures that were sourced from Gutiérrez et al., were based on the total costs.[136, 137] However, Gutiérrez et al. also provide the incremental costs of patients with fractures relative to matched controls. Since the incremental costs are more specific for the costs that are associated with the fracture and the model does not include additional costs of patients who do not sustain fractures beyond the disease management costs, the ERG considers it more appropriate to use the incremental costs for their base-case analysis. A similar approach based on incremental costs was also used in TA464 and ID901.[11, 87] The incremental first year costs provided by Gutiérrez et al., updated to 2019/2020 using the NHSCII as provided by the PSSRU 2020,[133] are £5,369 for a hip fracture, £1,465 for a vertebral fracture, and £877 for a NHNV fracture. A disadvantage of using these incremental cost estimates is that these do not include rehabilitation costs, which were included in the total cost for hip fracture used in the company's analyses.

The ERG notes that in TA464 a unit cost for long-term care was used and that was based on the assumptions that 1) equal proportions of patients who are discharged to long-term care go to nursing homes and residential care homes, 2) costs in the private sector are applicable (i.e., since the private sector provides 78% of places), and 3) that 36% of care is self-funded.[11] Using the unit costs as provided in PSSRU 2020,[133] £836 per week for private sector nursing homes and £620 per week for private sector residential care, the daily cost of long-term care can be estimated as 0.64 x (620+836) / 2 / 7 = £67. The ERG preferred to use this value for their base-case analysis.

4.2.9.5 Adverse event cost

Adverse event costs were applied to GIAEs at £40, based on a combination of the unit cost for a physician visit (see above) and a course of proton pump inhibitors (generic ranitidine, 300 mg tablets) at £0.90, sourced from the British National Formulary (BNF) January 2021.[131]

The company included the option to include CV AEs for those patients without a contraindicating history. The company identified the direct costs of CV events from a SLR from 2018.[139] This study estimated hospitalisation costs, outpatient referrals, primary care visits and medications of MI, stroke, unstable angina, heart failure, transient ischemic attack, and coronary artery bypass graft/percutaneous transluminal coronary angioplasty (CABG/PTCA), using hospital episodes statistics (HES) and CPRD data.[140] The estimated mean costs in the first 6 months after the first CV event was £4,594.16 in 2014 prices (£4993.85 in 2020, inflated using the indexes in Table 63 of the response to request for clarification[9]). Mean annualised cost in month 7 to 36 was £2,262.92 in 2014 prices (inflated to £2,459.79 in 2020 prices). The economic model was built to accommodate first and subsequent year costs, respectively. Therefore, the month 1 to 6 costs were applied in the first year and the month 7 to 36 costs were applied annually in every subsequent year until end of model time horizon or death. The company noted that this is likely to be a conservative approach as the first-year cost may be slightly overestimated in the model, since the majority costs likely occur closely to the event.[141]

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

Table 5.1 shows the deterministic CE results of the company's base-case analysis. All results are discounted and based on the confidential PAS price for romosozumab. Given that there are two relevant comparators, results are reported in a full incremental way. Pairwise ICERs of ROMO/ALN vs. each of the comparators (ALN and no treatment) are also reported for completeness. Results indicated that no treatment is dominated by ALN. Compared to ALN, ROMO/ALN accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER was £16,660 per QALY gained.

Table 5.1: Company base-case deterministic cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				3,747
ALN	[redacted]	10.014	[redacted]	[redacted]	0.021	[redacted]		16,660
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.031	[redacted]	16,660	

Based on Table 38 of the CS.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN

ALN = alendronate; CS = company submission; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = patient access scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

The disaggregated discounted costs are shown in Table 5.2 for the comparison vs. ALN and in Table 5.3 for the comparison vs. no treatment.

Table 5.2: Disaggregated cost results (ROMO/ALN vs. ALN)

Cost item	Cost intervention (ROMO/ALN)	Cost comparator (ALN)	Increment	Absolute increment	Absolute increment (%)
Hospitalisation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nursing home	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 1st treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 2nd treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 165 of CS Appendix J.[8]

ALN = alendronate; CS = company submission; ROMO/ALN = romosozumab-to-alendronate

Table 5.3: Disaggregated cost results (ROMO/ALN vs. no treatment)

Cost item	Cost intervention (ROMO/ALN)	Cost comparator (no treatment)	Increment	Absolute increment	Absolute increment (%)
Hospitalisation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nursing home	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 1st treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 2nd treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 166 of CS Appendix J.[8]
ALN = alendronate; CS = company submission; ROMO/ALN = romosozumab-to-alendronate

The company did not present disaggregated results for QALYs but reported differences in fracture events over 10 years between treatment arms, which is the main driver of the difference in QALYs produced by the model. These results are displayed in Tables 5.4 and 5.5.

Table 5.4: Summary of number of fracture events over 10 years, ROMO/ALN vs. ALN

Fracture type	Fracture events intervention (ROMO/ALN)	Fracture events comparator (ALN)	Difference
Hip	[redacted]	[redacted]	[redacted]
Vertebral	[redacted]	[redacted]	[redacted]
NHNV	[redacted]	[redacted]	[redacted]
Any	[redacted]	[redacted]	[redacted]

Based on Table 163 of CS Appendix J.[8]
ALN = alendronate; CS = company submission; NHNV = non-hip, non-vertebral; ROMO/ALN = romosozumab-to-alendronate

Table 5.5: Summary of number of fracture events over 10 years, ROMO/ALN vs. no treatment

Fracture type	Fracture events intervention (ROMO/ALN)	Fracture events comparator (no treatment)	Difference
Hip	[redacted]	[redacted]	[redacted]
Vertebral	[redacted]	[redacted]	[redacted]
NHNV	[redacted]	[redacted]	[redacted]
Any	[redacted]	[redacted]	[redacted]

Source: Table 164 in CS Appendix J.[8]
ALN = alendronate; CS = company submission; NHNV = non-hip, non-vertebral; ROMO/ALN = romosozumab-to-alendronate.

Overall, the technology is modelled to affect QALYs by:

- Reducing the incidence of fractures.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments, and
- Reducing costs associated to a decreased number of fractures.

5.2 Company's sensitivity and scenario analyses

5.2.1 Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) in which all input parameters were sampled simultaneously from their corresponding probability distributions over 1,000 iterations. The input parameters and the probability distributions used in the PSA can be seen in Table 36 of the CS.[1] The main distributional assumptions for the model parameters highlighted by the company are described below:

- Drug unit costs are assumed to be fixed and, therefore, they are not sampled in the model. For all the other cost parameters, a lognormal distribution with a standard error of 25% of the base-case value was assumed.
- Utility multipliers for hip, vertebral and NHNV fractures were sampled from a lognormal distribution with standard errors based on study data.
- Persistence on treatment and proportions of patients going to long-term care after a hip fracture were sampled from a beta distribution.
- Risk ratios for treatment efficacy were sampled from a normal distribution. Standard errors were based on the trial data and/or NMA.

The average PSA results are summarised in Table 5.6, and presented on a CE plane in Figure 5.1, from which a CE acceptability curve (CEAC) was calculated and plot in Figure 5.2. Both the CE-plane and CEAC plots are based on the pairwise comparisons vs. ROMO/ALN.

Table 5.6: Company base-case probabilistic cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER** (£/QALY)
No treatment	[redacted]	NR	[redacted]	Dominated by ALN				3,976*
ALN	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]		14,537
ROMO/ALN	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]	14,537	

Based on Table 39 of the CS.[1]

* Not the same as in the CS, probably due to rounding of QALYs; ** All pairwise ICERs are calculated vs. ROMO/ALN

ALN = alendronate; CS = company submission; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; NR = not reported; PAS = patient access scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

The average PSA results are in line with the deterministic ones shown in Table 5.1. Also, in the PSA no treatment is dominated by ALN, and the ICER for the comparison ROMO/ALN vs. ALN was £14,537 per QALY gained. The lower PSA ICER is the result of both lower incremental costs and higher incremental QALYs for ROMO/ALN vs. ALN. As shown in Figure 5.1, at the threshold of

£30,000 per QALY gained, the estimated probability that ROMO/ALN is a cost-effective alternative to ALN was [redacted] and [redacted] compared to no treatment.

Figure 5.1: Probabilistic sensitivity analysis cost effectiveness plane (PAS price for romosozumab)

Based on Figure 16 of the CS.[1]
Note: mind the axes of the CE-plane; they are not presented in their most common form (x-axis for incremental QALYs and y-axis for incremental costs)
ALE = alendronate; CE = cost effectiveness; CS = company submission; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROM = romosozumab

Figure 5.2: Probabilistic sensitivity analysis cost effectiveness acceptability curve (PAS price for romosozumab)

Based on Figure 17 of the CS.[1]
ALE = alendronate; CS = company submission; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROM = romosozumab; WTP = willingness to pay

5.2.2 Deterministic sensitivity analysis

The company also conducted deterministic sensitivity analyses (DSAs) where key parameters were individually varied at lower and upper bounds of values that were deemed plausible by the company. These are summarised in Table 5.7. Note that parameters like the starting age in the model, the length of the time horizon or the duration of the offset time are usually not included in the DSA but in scenario analyses.

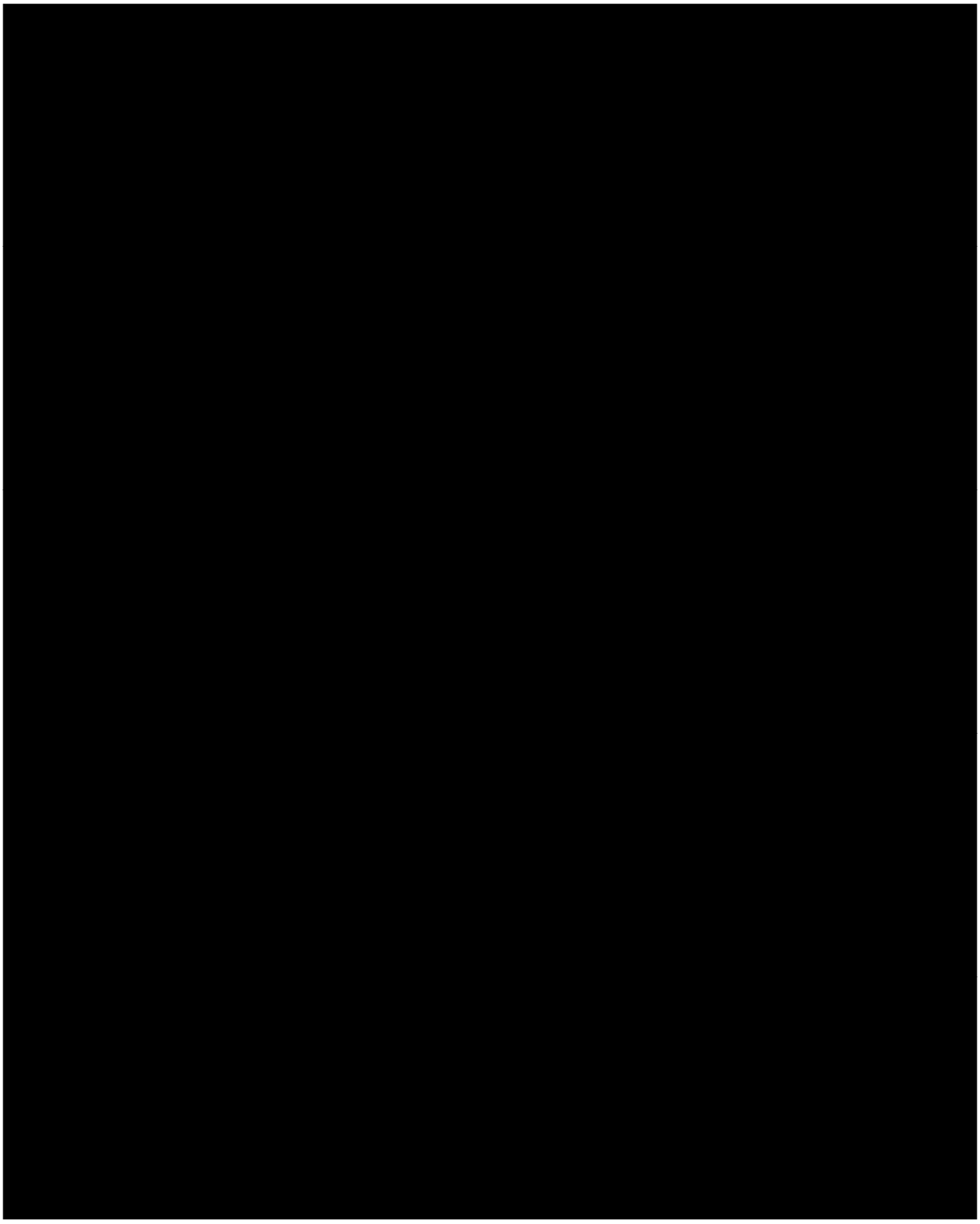
Table 5.7: Parameters and values included in the company's DSA

Parameter	Values	ERG comment
Start age	50, 60, 70 and 80 years	Scenario analyses (not DSA)
Model time frame	5, 10, 15 and 20 years	Scenario analyses (not DSA)
Fixed offset time	1, 2, 3, 4, 5 and 6 years	Scenario analyses (not DSA)
Utility multiplier for hip, vertebral and NHNV fracture in the first year following fracture	95% CI	Agree, evidence based
Utility multiplier for hip, vertebral and NHNV fracture in the second and following years after fracture	95% CI	Agree, evidence based
Direct medical cost first year after fracture	±25% of base-case	Agree, commonly used
Direct medical cost second and following years after hip and vertebral fracture	±25% of base-case	Agree, commonly used
Daily cost for long term care after hip fracture	±25% of base-case	Agree, commonly used
RRs for hip, vertebral and NHNV fractures for romosozumab	95% CI	Agree, evidence based
RRs for hip, vertebral and NHNV fractures for alendronate	95% CI	Agree, evidence based
Persistence multiplier for romosozumab	±25% of base-case	Arbitrary
Persistence multiplier for alendronate	±25% of base-case	Arbitrary

Based on Table 40 in CS.[1]
CI = confidence interval; CS = company submission; DSA = deterministic sensitivity analysis; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; NHNV = non-hip, non-vertebral; RR = relative risk

The results of the DSAs are presented in Table 41 in the CS.[1] This table shows pairwise ICERs for the comparisons ROMO/ALN vs. ALN and ROMO/ALN vs. no treatment for all scenarios defined in Table 5.7. For the comparison vs. no treatment, all ICERs were below the £20,000 per QALY gained threshold (or ROMO/ALN was dominant), except for the following scenarios: start age 50 years (ICER was £28,721), start age 60 years (ICER was £31,642) and time horizon 5 years (ICER was £49,862). The ICER was more sensitive to changes for the comparison vs. ALN. The results for this comparison were summarised by the company in the form of a tornado diagram as shown in Figure 5.3. This shows that the model results are sensitive to varying the time horizon, persistence, start age, changes in treatment effect of romosozumab/alendronate and alendronate alone, and utility multipliers for hip, vertebral and NHNV fracture.

Figure 5.3: DSA tornado diagram for romosozumab/alendronate vs. alendronate (PAS price for romosozumab)



Based on Figure 18 of the CS.[1]
CS = company submission; DSA = deterministic sensitivity analysis; GI = gastrointestinal; ICER = incremental cost effectiveness ratio; NHNV = non-hip, non-vertebral; QALY = quality-adjusted life year; RR = relative risk; vert = vertebral

5.2.3 Scenario analysis

The company conducted several scenario analyses in which the CE of ROMO/ALN was analysed against comparators that were not included in the base-case analysis. A summary of these scenarios is provided in Table 5.8. Scenario analyses 1 to 9 were based on the NMA using the ITT populations of ARCH and FRAME. Scenario 10 was based on the NMA using the EU label-matched populations from ARCH and FRAME. A patient population with a recent MOF, 74 years at treatment start, T-score of −2.9 and fracture risk corresponding to approximately 30% based on FRAX was assumed for scenarios 1 to 10. Scenario 11 was conducted for the comparison of ROMO/ALN vs. denosumab, as in scenario 7, but assuming a patient population at a higher risk of fracture. In particular, the assumed patient population for this scenario consisted of 74-year-old women, with a recent MOF and a T-score of −[redacted] and an approximately 10-year probability of MOF of [redacted]% according to FRAX. The results of the scenario analysis are presented in Table 5.9. All results include PAS price for ROMO. Results showed that ROMO/ALN was dominant or ICERs below £20,000 per QALY gained except for the comparisons against denosumab in the base-case population (£35,400 in scenario 7) and in the higher risk population (£27,509 in scenario 11).

Table 5.8: Summary of company scenario analyses

Scenario	Comparison	Treatment length	Offset	NMA efficacy source
1	ROMO/ALN vs. ALN	ROMO: 12m ALN: 48m vs. ALN: 60m	Dynamic	ITT population
2	ROMO/ALN vs. TERI	ROMO: 12m ALN: 48m vs. TERI: 24m	Dynamic	ITT population
3	ROMO/ALN vs. TERI	ROMO: 12m ALN: 48m vs. TERI: 18m	Dynamic	ITT population
4	ROMO/ALN vs. TERI biosimilar/ALN	ROMO: 12m ALN: 48m vs. TERI bio: 18m ALN: 42m	Dynamic	ITT population
5	ROMO/ALN vs. TERI/ALN	ROMO: 12m ALN: 48m vs. TERI: 18m ALN: 42m	Dynamic	ITT population
6	ROMO/ALN vs. RAL	ROMO: 12m ALN: 48m vs. RAL: 60m	Dynamic	ITT population
7	ROMO/ALN vs. DENO	ROMO: 12m ALN: 48m vs. DENO: 60m	ROMO: Dynamic DENO: 12m	ITT population
8	ROMO/ALN vs. RIS	ROMO: 12m ALN: 48m vs. RIS: 60m	Dynamic	ITT population
9	ROMO/ALN vs. ZOLE	ROMO: 12m ALN: 48m vs. ZOLE: 60m	Dynamic	ITT population
10	ROMO/ALN vs. ALN	ROMO: 12m ALN: 48m vs. ALN: 60m	Dynamic	ARCH EU*
11	ROMO/ ALN vs. DENO**	ROMO: 12m ALN: 48m vs. DENO: 60m	ROMO: Dynamic DENO: 12m	ITT population

Source: Table 42 and 43 in CS.[1]

* ARCH-EU label-matched population used in NMA. ** Scenario conducted for a population with a higher risk of fracture.

Note: For DENO, the company assumed a clinical effect limited to within 6 months after stopping treatment.[27, 111] The company explained that chronic treatment with DENO is necessary when used as the subsequent treatment after ROMO for this combination to provide optimal benefits to patients; or alternatively a further treatment switch to a bisphosphonate after the DENO treatment period would be required. Therefore, a 1-year fixed offset time was applied to DENO.

ALN = alendronate; CS = company submission; DENO = denosumab; EU = European Union; ITT = intention-to-treat; m = months; NMA = network meta-analysis; RAL = raloxifene; RIS = risedronate, ROMO = romosozumab; TERI = teriparatide, ZOLE = zoledronate

Table 5.9: Company scenario analyses results (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER (£/QALY)*
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	9.993	[redacted]	Dominated by DENO				3,747
RALO (6)	[redacted]	9.998	[redacted]	Dominated by DENO				Dominated
RIS (8)	[redacted]	10.013	[redacted]	Dominated by DENO				12,518
ALN (1)	[redacted]	10.014	[redacted]	Dominated by DENO				16,660
TERI (3)	[redacted]	10.021	[redacted]	Dominated by DENO				Dominated
TERI (2)	[redacted]	10.023	[redacted]	Dominated by DENO				Dominated
TERI/ALN (5)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
TERI biosimilar/ALN (4)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
ZOLE (9)	[redacted]	10.026	[redacted]	Dominated by DENO				17,176
DENO (7)	[redacted]	10.034	[redacted]					35,400
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.011	[redacted]	35,400	
Scenario 10								
ALN	[redacted]	10.013	[redacted]					
ROMO/ALN	[redacted]	10.043	[redacted]	[redacted]	0.030	[redacted]	17,690	
Scenario 11								
DENO	[redacted]	9.800	[redacted]					
ROMO/ALN	[redacted]	9.813	[redacted]	[redacted]	0.013	[redacted]	27,509	

Based on Tables 44, 45 and 46 of the CS.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; CS = company submission; DENO = denosumab; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = patient access scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

5.2.4 Additional scenario analysis requested by the ERG

Some areas of uncertainty were identified by the ERG during the clarification phase, resulting in the company conducting additional scenario analyses requested by the ERG in the clarification letter.[9] The uncertainties explored by the company in these additional scenarios are the following:

- **Imminent risk of fracture**. The ERG requested a scenario where the imminent risk of fracture was removed from the analysis. This was because the ERG considers it unclear whether the reduction in fracture risk from treatment, estimated from the ARCH ITT population, corresponds to a population with imminent risk of fracture (see Key issue 1). The company indicated that this scenario should not be considered relevant for this appraisal because it does not accurately represent the romosozumab target patient population. While this might be the case, given the uncertainties previously mentioned, the ERG considers that this scenario provides relevant information. Results are shown in Table 5.10. The ICER increased by £18,523 compared to the base-case ICER of £16,660 per QALY gained.
- **Incidence of vertebral fractures**. Following the discussion in Section 4.2.6.1, the ERG asked the company to conduct a scenario analysis where the results from Singer et al. 1998 for vertebral fractures were assumed.[81] The company explained that the vertebral fracture incidences estimate from this study are generally not considered to be reliable. For that reason, the results from this scenario should not be considered relevant for this appraisal because, according to the company, it likely underestimates the risk of clinical vertebral fractures and therefore underestimates the CE of romosozumab. However, given the uncertainties concerning the company's approach described in Section 4.2.6.1, the ERG considers that this scenario has informational value, in only for providing an upper limit for the ICER with regards to the uncertainty about the incidence of vertebral fractures. Results are shown in Table 5.11. The ICER increased by £14,052 compared to the base-case ICER of £16,660 per QALY gained.
- **Treatment effect estimated from an alternative NMA**. The company considered that results for alendronate vs. placebo were similar in both the NICE and the CS NMA. The ERG argued that this is a subjective statement seeing the values presented in Table 4.8, especially for the values shown for teriparatide. This raised concerns about the validity/credibility of the NMA results. Hence, the ERG asked the company to provide results based on the NICE NMA. The company concluded that CE scenarios based the NICE NMA are not appropriate for this appraisal because the underlying evidence base for such NMA was outside the licensed indication for romosozumab. However, given the uncertainties concerning the company's NMA, as highlighted in Key issue 3, the ERG considers this a valid scenario. The results in Table 5.12, show that the ICER was similar to the ICER in the company base-case.
- **Persistence**. Persistence assumptions were identified as one of the most important drivers of the CE results. Concerns regarding the company base-case assumptions on persistence and how these could bias the results in favour of romosozumab were explained in Section 4.2.6.2. Based on these, the ERG asked the company to explore three additional scenarios in which 1) persistence was assumed to be as in the ARCH trial for romosozumab and the alendronate, 2) persistence on romosozumab was assumed to be equal to persistence on teriparatide and 3) an unrealistic scenario with 100% compliance in both intervention and comparator. Again, the company indicated that these scenarios are not relevant for this appraisal. In particular, for the first scenario, the company emphasised that persistence inputs derived from clinical trials are known to differ substantially from real-world persistence of osteoporosis patients and are at high risk to misrepresent the CE of romosozumab. The ERG agrees with this and as explained in Section 4.2.6.2, considers that by using trial-based persistence for romosozumab vs. real life

persistence for alendronate, there is indeed a high risk that the CE of romosozumab is misrepresented in the company base-case. Even though it is known that real-life persistence will be lower than in trial settings, at least this scenario would provide a fair comparison. For the second scenario, the company considered that persistence to romosozumab is unlikely to be equal to teriparatide's persistence given the difference in administration frequency (romosozumab is given monthly and teriparatide is given daily). While the ERG acknowledged that this might be the case, the company has not provided evidence to support this assumption. Hence, the relevance of this scenario. Finally, even if it seems clear that a scenario based on 100% persistence is unrealistic, the results of this scenario can still be relevant for decision-making. Results are shown in Tables 5.13 to 5.15. In all scenarios the ICER increased compared to the base-case, especially in the first one where the ICER was almost £40,000 higher.

- **Alternative risk of death for hip and vertebral fractures**. The company run a scenario where the relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al. 2007 (UK setting).[142] The relative risks in the second and following years for hip and vertebral fractures, and first year for NHNV fractures, were assumed to be the same as in the base-case. Results are shown in Table 5.16. This had a minor impact on the CE results.
- **CV adverse events**. The ERG asked the company to include in the analysis CV AEs according to the incidence in the ARCH trial and relevant disutilities and costs. The company indicated that the results of this scenario can be considered conservative for romosozumab since the CV occurrence rates for romosozumab and alendronate were chosen from the study where the imbalance between these two treatments was greatest (ARCH) and subsequent year costs are applied every year after the CV event until the end of the modelled time horizon or death. The decision not to select or pool any other romosozumab studies (FRAME, STRUCTURE, McClung) where the CV event rate for romosozumab was lower than in ARCH to derive CE results of this scenario means that the results should be considered to be extremely conservative, and for illustrative purposes only. Nonetheless, the ERG considers that since the efficacy results are based on ARCH it is appropriate that AE evidence is based on ARCH. Results are shown in Table 5.17. The ICER increased by £2,840 compared to the base-case.
- **Drug administration costs**. The company ran a scenario including drug administration costs (i.e., for subcutaneous injections) when the PSP is not in place for romosozumab, as well as for the relevant comparators that are used in scenario analyses. The cost (£9.5 per administration) was based on a 15-minute visit (based on £38 per hour for GP nurse contact time). PSSRU Unit Costs of Health and Social Care 2020 10.2 Nurse (GP practice). Unit costs available 2019/2020 based on 1,573 hours per year, which includes 225 working days minus sickness absence (8 days) and any training/study days as reported for all NHS staff groups. In the scenario analysis, romosozumab is associated with 12 SC injections days (i.e., 24 injections) per year administered by a nurse; teriparatide 365 injections per year and denosumab two injections per year. Results are shown in Table 5.18. All ICERs increased (moderately) compared to those shown in Table 5.9.

Table 5.10: Company scenario with fracture recency removed (no imminent risk) cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	10.044	[redacted]	Dominated by ALN				12,688

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ALN	[redacted]	10.055	[redacted]	[redacted]	0.011	[redacted]		35,183
ROMO/ALN	[redacted]	10.074	[redacted]	[redacted]	0.019	[redacted]	35,183	

Based on Table 44 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.11: Company scenario with vertebral fracture incidences from Singer et al. 1998 cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	10.069	[redacted]	Dominated by ALN				8,967
ALN	[redacted]	10.075	[redacted]	[redacted]	0.006	[redacted]		30,712
ROMO/ALN	[redacted]	10.087	[redacted]	[redacted]	0.012	[redacted]	30,712	

Source: Based on Table 45 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate.

Table 5.12: Company scenario using efficacy of ALN vs. placebo from NICE NMA cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				4,219
ALN	[redacted]	10.013	[redacted]	[redacted]	0.020	[redacted]		17,069
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.032	[redacted]	17,069	

Based on Table 47 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; NMA = network meta-analysis; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.13: Company scenario with persistence data based on ARCH for all treatments cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				646

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ALN	[redacted]	10.034	[redacted]	[redacted]	0.041	[redacted]		54,340
ROMO/ALN	[redacted]	10.051	[redacted]	[redacted]	0.017	[redacted]	54,340	

Based on Table 53 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.14: Company scenario with romosozumab persistence equal to teriparatide persistence cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				10,016
ALN	[redacted]	10.014	[redacted]	[redacted]	0.021	[redacted]		38,295
ROMO/ALN	[redacted]	10.032	[redacted]	[redacted]	0.018	[redacted]	38,295	

Based on Table 54 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate.

Table 5.15: Company scenario with 100% persistence for all treatments cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN and ROMO/ALN				Dominated
ALN	[redacted]	10.045	[redacted]	[redacted]	0.052	[redacted]		20,989
ROMO/ALN	[redacted]	10.072	[redacted]	[redacted]	0.027	[redacted]	20,989	

Based on Table 55 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.16: Company scenario using relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al. 2007 cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.981	[redacted]	Dominated by ALN				3,824
ALN	[redacted]	10.000	[redacted]	[redacted]	0.019	[redacted]		16,728

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ROMO/ALN	[redacted]	10.031	[redacted]	[redacted]	0.031	[redacted]	16,728	

Based on Table 59 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.17: Company scenario including cardiovascular adverse events cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.966	[redacted]	Dominated by ALN				5,075
ALN	[redacted]	9.986	[redacted]	[redacted]	0.020	[redacted]		19,500
ROMO/ALN	[redacted]	10.013	[redacted]	[redacted]	0.027	[redacted]	19,500	

Based on Table 60 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.18: Company scenario analyses results including cost for subcutaneous administrations (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	9.993	[redacted]	Dominated by DENO				5,123
RAL (6)	[redacted]	9.998	[redacted]	Dominated by DENO				Dominated
RIS (8)	[redacted]	10.013	[redacted]	Dominated by DENO				14,953
ALN (1)	[redacted]	10.014	[redacted]	Dominated by DENO				19,434
TERI (3)	[redacted]	10.021	[redacted]	Dominated by DENO				Dominated
TERI (2)	[redacted]	10.023	[redacted]	Dominated by DENO				Dominated
TERI/ALN (5)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
TERI biosimilar/ALN (4)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
ZOLE (9)	[redacted]	10.026	[redacted]	Dominated by DENO				21,129
DENO (7)	[redacted]	10.034	[redacted]					43,000
ROMO/ALN	[redacted]	10.044	[redacted]	[redacted]	0.010	[redacted]	43,000	

Source: Based on Table 61 of the clarification letter response.[9]
Note: It is unclear why Table 61 of the clarification letter response provides different QALYs/LYG than those in Table 5.10 since only costs are supposed to change.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; DENO = denosumab; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RAL = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

5.2.5 Conclusions from company's sensitivity and scenario analyses

The modelling assumptions that have the greatest effect on the ICER are:

- Treatment persistence.
- Start age of the population.
- Model time horizon.
- Treatment effect of romosozumab followed by alendronate and alendronate alone.
- Utility multipliers for hip, vertebral and NHNV fracture.
- Comparator choice (denosumab).

5.3 Model validation and face validity check

Validation efforts conducted on the economic model were discussed in the validation section of the CS (B.3.10).[1] In general, the model adheres to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] A comparison between the recommended key modelling aspects and the assumption made on the romosozumab model is provided in Table 47 of the CS.[1] Modelling assumptions were also validated by leading UK experts at an advisory board held by the company in 2017.

Most of the validation efforts discussed in the CS referred to those conducted by NICE's PRIMA (Preliminary Independent Model Advice service) in 2017.[122, 143] PRIMA assessed the appropriateness of the conceptual model, model verification (through black-box testing), reproducibility and made suggestions on how to improve the model's transparency and usability. The complete PRIMA report was presented as part of the CS. Furthermore, the company explained that the model has also undergone thorough quality control by Quantify Research, including performing multiple verification and validation tests, as well cross-validating the results with another in-house osteoporosis model.

The company also mentioned that the same model has been used in two published peer-reviewed manuscripts,[73, 74] and in the reimbursement submissions of romosozumab in Sweden (TLV) and

Scotland (SMC).[104, 125] Although not explicitly stated, it is assumed that the model might have also passed quality controls previous to publication and/or during the reimbursement assessments.

Additionally, the number of fractures predicted by the CE model was validated using a Swedish cohort study of women 50 years and older with fracture identified in the National Patient Register. Details of the Swedish registry study can be found elsewhere.[86] Since BMD data were available from three large hospitals in Sweden, a comparison between the model and real-world fracture incidences adjusted for risk factors such as age and BMD was possible. For this comparison, the romosozumab model was populated with Swedish population incidences and used the Swedish version of FRAX. Using the registry data, the incidence of fracture (all types) was predicted for 5-year follow-up with a multiple-failure model. The 10-year incidence was calculated using the non-parametric single-failure model. These were compared with the incidence predicted by the health economic model. The results of this comparison can be seen in Table 5.19. The CE model predicted approximately [redacted]% higher 5-year incidence than the incidence estimated from the registry data. The company considered that this can be is mainly explained by the fact that vertebral fractures are at risk of being underreported in register data. Ten-year incidence was calculated using register data for women 55 to 90 years with MOF and unknown BMD. However, the same population cannot be completely reproduced in the CE model, which makes this comparison of limited value. In the CE model, the fracture risk is likely to be higher than the fracture risk for the average Swedish population 55 to 90 years with unknown BMD. This is shown in Table 5.20. However, the extent to which the 10-year risk predicted by the model are comparable to the risk observed in real life is unknown.

Table 5.19: Validation of simulated fracture risks using Swedish register data

Source	Outcome	Women with MOF**, age 74, unknown BMD	Women with MOF**, age 74, T-score -2.9	Women with MOF, age 55-90**, unknown BMD
Register study	5-year cumulative incidence of new fracture (disregarding type)*	34.6% (1a)	52.5% (1b)	
CE model***	5-year cumulative incidence of new fracture	[redacted]% (1a)	[redacted]% (1b)	
Register study	10-year non-parametric cumulative incidence of a new fracture (single failure model)			37.6%
CE model***	10-year risk of a new fracture (single failure model)	[redacted]%	[redacted]%	

Based on Table 48 of the CS.[1]

* Predicted incidence based on a multiple failure model; ** At baseline; *** Excess mortality of fracture set to 100%. The CE model adjusts mortality for comorbidities, i.e., mortality unrelated to the fracture. This adjustment cannot be made in the register data; therefore, excess mortality was set to 100% in the model for better comparison.

BMD = bone mineral density, CE = cost effectiveness, CS = company submission; MOF = major osteoporotic fracture

Finally, in response to clarification question B27,[9] the company provided a comparison of the distribution of fractures in the Swedish real-world study vs. the distribution of fractures in the CE model. In the Swedish real-world study, out of the 231,769 patients with at least one fracture, 7,656 patients (3.3%) had a third fracture over approximately 5.5 years of maximum follow-up data.[85] The CE model estimated 4.4% of patients had a third fracture over 5 years. The company explained that these values are not strictly comparable since in the Swedish data, the first fracture could have happened at some point during the 5.5 years of follow-up, meaning that not all patients would have enough follow-up time to have developed a second or a third fracture.

ERG comment: The model adheres in general to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] Since 2017, the model has been involved in several iterations of quality assessment including the NICE PRIMA. In line with this assessment, the ERG considers that review would be facilitated if calculations were performed in the model work sheets, instead of being hard coded in VBA. As explained in Section 4.2.2, the VBA code was initially password protected because the FRAX algorithm is confidential. After clarification, the company provided the rest of the VBA which was reviewed by the ERG. The VBA code was well structured and sufficient comments were provided to understand the flow of the code. In reviewing the model and the VBA code, the ERG noted the following issues:

- In the 'State trace' sheet of the model the proportions of patients with a first NHNV fracture (i.e. column M) always remains zero, whereas from the second cycle onwards there is a non-zero proportion of patients with a second NHNV fracture. The ERG could not trace the source of this issue.
- After running the model with the 'Trackers summary' enabled the ERG noted that the means of outpatient costs do not match with the means of outpatient cost on the 'Results' sheet. From scrutinizing the VBA code in module mRunModel.bas, it appears that t_IterCost (comparator, 3) is not updated (lines 3264-3270) for costs in year 2 and more after hip and vertebral fracture. If this is indeed the cause, it seems that it does not impact the overall results.
- Also, in the 'Trackers summary' the drug costs and treatment management costs always remain zero but not in the 'Results' sheet. The ERG could not trace the cause of this. Note that the means of other costs, LYs and QALYs did match between the 'Trackers summary' and the 'Results' sheet.
- In the module mRunModel.bas an error was found in line 2065. In the formula PrevFx = PrevFx + t_fx(comparator, 1) + t_fx(comparator, 1) + t_fx(comparator, 3) the second 't_fx(comparator, 1)' should read 't_fx(comparator, 2)'. It is not clear to the ERG to what extent this impacts the results.

An additional point the ERG would like to emphasise is the model running time. Despite the added complexity of microsimulation compared to standard cohort models, the model seems to be extremely demanding regarding the computational power needed to run within reasonable time. Even a deterministic run would take more than 20 minutes. This makes the validation process extra difficult and for this reason, the ERG was not able to validate the results of some of the scenarios presented by the company. In particular, the ERG did not succeed in running any PSA. Sometimes the model would stop running after a few PSA iterations and most of the times Excel would crash. The default settings of 500,000 iterations for the inner loop and 1,000 for the outer loop projected a running time of more than 2 weeks to finish, which in practice can be deemed as unfeasible. Given this practical issue, the ERG would like to suggest the company to conduct an analysis to estimate the minimal PSA loop sizes that would provide reliable results in a minimum running time and to re-consider the programming of the model in order to make it computationally more efficient.

As explained in detail in the ERG comment in Section 4.2.6 (baseline fracture incidence), there is uncertainty regarding the validity of the incidences of hip, vertebral and NHNV fractures, relating to the aspects:

- The company used a study that dates from 1998 by Singer et al.[81] as the main source of input values.
- The company referred to a study by van der Velde et al. 2016 to confirm the stability of hip fracture incidence over time but which had substantially lower incidence rates than Singer et al. 1998.[81, 82]
- The company referred to a study by Kanis et al. 2001 to confirm the similarity between ratios of vertebral to hip fractures in Sweden and the UK.[90] The ERG could not confirm that a comparison between ratios of vertebral to hip fractures in Sweden and the UK was included in Kanis et al. 2001.
- For the different types of fractures that were included in the estimates of the incidence of NHNV fractures that were sourced from Singer et al. 1998, the company referred to van der Velde et al. 2016 to confirm the stability over time and similarity of findings from both studies.[81, 82] However, the ERG could not confirm the stability over time and the similarity of findings for all types of fractures that were included.

Validation was presented against Swedish data only. The company indicated that it was not possible to perform the validation based on UK data, since detailed data on fractures and risk factors such as BMD were not available. Therefore, it is uncertain whether the validity results can be generalised to the UK.

Comparisons with other TAs were not presented. Therefore, it is not possible to quantify whether the results in the CS are in line with those in previous appraisals.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

6.1.1 Explanation of the company adjustments after the request for clarification

During the clarification phase, the ERG requested the company to correct errors on the 'PSA input' sheet that resulted in cells displaying '#N/A' and '#NUM!'. The company provided a corrected version of the model alongside their response to the ERG's clarification questions.

6.1.2 Explanation of the ERG adjustments

The changes that the ERG can make (to the model received with the response to the clarification letter) can be subdivided into the following three categories (according to Kaltenthaler et al. 2016[144]):

- Fixing errors (correcting the model where the company's electronic model is unequivocally wrong).
- Fixing violations (correcting the model where the ERG considers that the NICE reference case, scope or best practice has not been adhered to).
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred).

In the current assessment, only matters of judgement played a role. After the proposed changes were implemented in the company's model, additional scenario analyses were explored by the ERG in order to assess the impact of alternative assumptions on the CE results.

6.1.2.1 Fixing errors

No errors were corrected by the ERG in the model provided in response to the clarification letter. Note that the ERG was granted access to a version of the model in which the VBA code was unprotected to facilitate validation by the ERG. However, the company was unable to perform exhaustive quality assurance on the "unprotected" version of the model and asked the ERG to use the model received with the response to the clarification letter to conduct all ERG scenarios. As a consequence, the ERG was not able to change any of the model VBA code, regardless of whether this was with the purpose of fixing errors or testing alternative assumptions.

6.1.2.2 Fixing violations

No violations were applicable to this appraisal.

6.1.2.3 Matters of judgement

The ERG's preferences regarding alternative assumptions led to the following changes to the company base-case analysis:

- Romosozumab persistence (i.e., at 6 and 12 months) was changed from 90% to 80% (see Section 4.2.6.2).
- Alendronate persistence was changed as follows: for alendronate after romosozumab the ERG used estimates for persistence with oral bisphosphonates in non-naïve patients from Morley et al. 2020 and for alendronate alone the ERG used estimates for persistence with oral bisphosphonates in naïve patients from Morley et al. 2020 (see Section 4.2.6.2).[100]
- Only excess mortality for hip fractures (and not for other types of fractures) was included in the analysis (see Section 4.2.6.4).

- Daily costs of long-term care were changed from £112 to £67 (see Section 4.2.9).
- The ERG changed the input parameter values for the costs associated with fractures from £13,203 to £5,369 for hip fractures, from £2,897 to £1,465 for vertebral fractures, and from £2,131 to £877 (see Section 4.2.9).
- Cardiovascular events which occurred in patients who did not have a history of MI or stroke were included in the analysis (see Section 4.2.7).
- Costs for administration of romosozumab (and for the comparators denosumab and teriparatide) that are applicable as long as the PSP is not in place were included in the analysis (see Section 4.2.9).
- The frequency of physician visits was changed from once per year to twice per year (see Section 4.2.9).
- General population mortality input parameter values were updated to the most recent UK National Life Tables (see Section 4.2.6.4).

The overview of the changes and the bookmarks for the justification of the ERG changes are presented in Table 6.1.

Table 6.1: Company and ERG base-case preferred assumptions

Base-case preferred assumptions		Company	ERG	Justification for change
Persistence with romosozumab		90%	80%	Section 4.2.6.2
Persistence with alendronate	**Alendronate after romosozumab**	85% of persistence with denosumab as reported in Karlsson et al. 2015[97]	Morley et al. 2020 persistence with oral BPs in non-naïve patients[100]	Section 4.2.6.2
	Alendronate alone	Li et al. 2012[96]	Morley et al. 2020 persistence with oral BPs in naïve patients[100]	
Excess mortality following fractures		Included for hip, vertebral and NHNV fractures	Included for hip fractures only	Section 4.2.6.4
Daily costs of long-term care		£112	£67	Section 4.2.9
Costs associated with fractures	**Hip**	£13,203	£5,369	Section 4.2.9
	Vertebral	£2,897	£1,465	
	NHNV	£2,131	£877	
Cardiovascular events		Not included	Included	Section 4.2.7
Romosozumab administration costs (PSP)		Not included (PSP in place)	Included (PSP not in place)	Section 4.2.9
Frequency of physician visits		Once per year	Twice per year	Section 4.2.9
General population mortality		2012-2014 UK National Life Tables	2017-2019 UK National Life Tables	Section 4.2.6.4

BP = bisphosphonates; ERG = Evidence Review Group; NHNV = non-hip, non-vertebral; PSP = Patient Support Programme; UK = United Kingdom

6.1.3 Additional scenarios conducted by the ERG

The ERG conducted a series of scenario analyses to explore the impact of key assumptions and uncertainties within the CE analyses. These uncertainties were related to the inclusion of comparators other than alendronate alone, removal of the imminent risk, persistence, the PSP, costs associated with fractures, utilities, AEs, treatment effect waning, and excess mortality associated to fractures.

6.1.3.1 Scenario set 1: other comparators

The ERG performed scenario analyses using the same comparators defined by the company in Section 5.2.3: teriparatide, zoledronate, denosumab, risedronate, raloxifene and no treatment.

6.1.3.2 Scenario set 2: imminent risk removed

To address the uncertainty regarding the relevant population for this appraisal, as discussed for example in Section 4.2.3, the ERG performed a set of scenario analyses where the "imminent risk" of fracture was removed from the analysis. This set of scenarios was performed with all comparators as in scenario set 1.

6.1.3.3 Scenario set 3: persistence

To address the uncertainty regarding assumptions on persistence with osteoporosis therapies, the ERG performed the following set of scenario analyses:

- No distinction is made between alendronate naïve (i.e., patients receiving alendronate alone) and non-naïve patients (i.e., patients receiving alendronate after romosozumab). Thus, this scenario assumes the same persistence for patients receiving alendronate after romosozumab and alendronate alone, both persistence estimates based on persistence with oral BPs in Morley et al. 2020 in the 'All patients' (i.e., naïve patients and non-naïve patients pooled) population.[100]
- An analysis where it is assumed that persistence with romosozumab is the same as in the company base-case; i.e., 90% instead of 80%.
- A scenario was also conducted assuming persistence for romosozumab as per the ERG base-case and persistence for alendronate as per the company base-case.
- The persistence scenarios requested at clarification were also repeated on the ERG base-case, including using the ARCH trial persistence for both romosozumab and alendronate; assuming the persistence on romosozumab was equal to that of teriparatide and assuming 100% persistence for all treatments.

6.1.3.4 Scenario set 4: patient support programme in place

To address the uncertainty regarding the impact on CE results following the implementation of the company's plans to set up the PSP, the ERG performed a set of scenario analyses where no administration costs are assumed for romosozumab and where the assumption of no administration costs is combined with the assumption of 90% persistence with romosozumab.

6.1.3.5 Scenario set 5: costs associated with fractures

To address the uncertainty regarding the costs associated with fractures, the ERG performed a scenario analysis assuming total health care costs associated with fractures from Gutiérrez et al. 2011 and 2012 (i.e., the same as in the company base-case analysis, which also includes rehabilitation costs for hip fractures), instead of the incremental costs of patients with fractures vs. those without (as in the ERG base-case analysis, which does not include rehabilitation costs) from the same sources.[136, 137]

6.1.3.6 Scenario set 6: utility multipliers

Although the application of utility multipliers for fracture events has been a common approach in previous osteoporosis appraisals[11, 87], the multipliers differ somewhat across appraisals. Therefore, scenarios using the alternative sets of multipliers (shown in Table 4.19 of this report) were conducted to examine the impact on results.

6.1.3.7 Scenario set 7: adverse events

The ERG included those CV AEs which occurred in patients without a history of MI or stroke in their base-case as an imbalance was observed in the ARCH trial. A scenario was also conducted where these CV AEs were excluded.

6.1.3.8 Scenario set 8: treatment effect waning

The ERG run a scenario in which 4 years of full treatment effect was assumed followed by a waning in effect for one more year. The fracture risk ratios assumed for the fifth year were the following: [redacted] for hip fracture, [redacted] for vertebral fracture and [redacted] for NHNV fractures. The dynamic offset was equal to 5 years.

6.1.3.9 Scenario set 9: excess mortality associated to fractures

Following ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] the ERG base-case included excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNV fractures were also explored by the ERG.

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.2.1 Results of the ERG preferred base-case scenario

The ERG preferred base-case incremental CE results, provided in Table 6.2, indicate that the total costs associated with romosozumab (12 months) followed by alendronate (48 months) were estimated at [redacted] and the total costs associated with alendronate alone (60 months) were estimated at [redacted], indicating an incremental cost of [redacted]. Total QALYs associated with romosozumab (12 months) followed by alendronate (48 months) were estimated at [redacted] and total QALYs associated with alendronate alone (60 months) were estimated at [redacted], indicating an incremental number of [redacted] QALYs gained. These results indicate an estimated ICER of £483,750 per QALY gained.

It should be highlighted that in the ERG base-case, the incremental LYGs are negative. This is due to the inclusion of serious CV AEs in the ERG base-case, which occurred more frequently in the romosozumab arm than in the alendronate alone arm, and which had an impact on mortality.

Table 6.2: ERG preferred base-case deterministic cost effectiveness results (discounted, PAS)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Romosozumab followed by alendronate	[redacted]	10.048	[redacted]	[redacted]	-0.002	[redacted]	483,750
Alendronate alone	[redacted]	10.050	[redacted]				

Based on the ERG preferred version of the electronic model.[1]
Note: The results of the comparison vs. no treatment are reported in Section 6.2.2.1 of the ERG report.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year.

As mentioned in Section 5.3, the ERG was unable to run a PSA for its preferred base-case analysis. However, given the deterministic ICER and assuming that the PSA ICER would be in line with this one, the probability that romosozumab is considered cost effective at a threshold of £30,000 compared to alendronate is likely to be [redacted]%.

6.2.2 Results of the ERG additional exploratory scenario analyses

6.2.2.1 Scenario set 1 results: other comparators

The results of scenario analyses set 1, using various alternative comparators, are provided in Table 6.3. These indicate that the relevant comparison is zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab followed by alendronate, show that all ICERs are above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 1 month, teriparatide 24 months and teriparatide followed by alendronate, which are dominated by romosozumab followed by alendronate; and the comparison against zoledronate, which is dominant.

Some counterintuitive results were observed when teriparatide was involved as a comparator treatment. It is unclear why the sequence teriparatide (or biosimilar) would result in less QALYs than teriparatide alone (even if teriparatide alone is given for 24 months and for 18 months as part of the sequence). If this would be the case, it would seem irrational to treat patients with the sequence when teriparatide alone is more beneficial. Also, note that this was not observed in the results presented by the company in Table 5.9. Therefore, the ERG explored this potential issue a bit further and run an "extreme" scenario in which teriparatide 18 months was compared with teriparatide 18 months followed by alendronate, but with persistence on alendronate equal to zero. In this scenario, teriparatide alone resulted in [redacted] QALYs and the sequence with alendronate at zero persistence resulted in [redacted] QALYs. Thus, the sequential treatment provided more QALYs even when persistence on the second treatment on the sequence was equal to zero. A similar scenario was run but with romosozumab instead of teriparatide and the same effect on QALYs was observed. The ERG was not able to find the source for these inconsistencies, which might need further confirmation from the company. It is also unclear why the sequence with teriparatide biosimilar would result in more QALYs than the sequence with commercial teriparatide. This is likely due to both sequences being informed by different NMAs.

Table 6.3: Scenario set 1 results: other comparators (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	10.0440	[redacted]	Dominated by RIS				£44,288
RALO (6)	[redacted]	10.0397	[redacted]	Dominated by RIS				£37,000
RIS (8)	[redacted]	10.0493	[redacted]	Dominated by ALN				£226,438
TERI (3) 18m	[redacted]	10.0509	[redacted]	Dominated by ALN				Dominated by ROMO/ALN
ALN (1)	[redacted]	10.0500	[redacted]					£483,750
DENO (7)	[redacted]	10.0532	[redacted]	Dominated by ZOLE				£1,088,000
TERI/ALN (5)	[redacted]	10.0516	[redacted]	Dominated by TERI bio/ALN				Dominated by ROMO/ALN
TERI biosimilar/ALN (4)	[redacted]	10.0516	[redacted]	Extendedly dominated by ROMO/ALN				£228,000
TERI (2) 24m	[redacted]	10.0515	[redacted]	Dominated by ROMO/ALN				Dominated by ROMO/ALN
ROMO/ALN	[redacted]	10.0484	[redacted]	Dominated by ZOLE				ZOLE dominates
ZOLE (9)	[redacted]	10.0492	[redacted]	[redacted]	-0.001	[redacted]	£47,583	

Based on the ERG preferred version of the electronic model.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; DENO = denosumab; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

6.2.2.2 Scenario set 2 results: imminent risk removed

The results of scenario analyses set 2, with the imminent risk removed, are provided in Table 6.4. These indicate that the relevant comparison is zoledronate vs. alendronate, with an ICER of £121,730 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab followed by alendronate, show that all ICERs are well above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 18 months, and teriparatide followed by alendronate, which are dominated by romosozumab followed by alendronate; and the comparisons against zoledronate and denosumab, which are dominant. The same counterintuitive results discussed in the previous section were also observed in this set of scenarios. Furthermore, it also seems counterintuitive that raloxifene was dominated by no treatment. However, this can be explained by looking at fracture risk ratios presented in Table 4.11. Therefore, the model results for this scenario seem consistent with the NMA input but the ERG is concerned about the validity of the value provided by the NMA.

Table 6.4: Scenario set 2 results: imminent risk removed

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
RALO (6)	[redacted]	10.0508	[redacted]	Dominated by no treatment				£76,548
No treatment	[redacted]	10.0543	[redacted]	Dominated by RIS				£98,965
RIS (8)	[redacted]	10.0591	[redacted]	Dominated by ALN				£667,218
TERI (3) 18m	[redacted]	10.0595	[redacted]	Dominated by TERI/ALN				Dominated by ROMO/ALN
TERI/ALN (5)	[redacted]	10.0601	[redacted]	Dominated by TERI bio/ALN				Dominated by ROMO/ALN
TERI bio/ALN (4)	[redacted]	10.0601	[redacted]	Dominated by ALN				£3,454,305
ROMO/ALN	[redacted]	10.0581	[redacted]	Dominated by ALN				
ALN (1)	[redacted]	10.0599	[redacted]					ALN dominates
TERI (2) 24m	[redacted]	10.0609	[redacted]	Dominated by DENO				£11,872,642
DENO (7)	[redacted]	10.0619	[redacted]	Dominated by ZOLE				DENO dominates
ZOLE (9)	[redacted]	10.0596	[redacted]	[redacted]	-0.0003	[redacted]	£121,730	ZOLE dominates

Based on the ERG preferred version of the electronic model.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; DENO = denosumab; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

6.2.2.3 Scenario set 3 results: persistence

The results of scenario analyses set 3, using various alternative assumptions and inputs for persistence, are provided in Table 6.5. These scenario results demonstrate the substantial and varied impact of different persistence assumptions on results. Using the persistence estimates from Morley et al. 2020,[100] based on all patients for persistence with oral BPs, substantially increased the incremental QALYs and reduced the ICER by approximately £400,000 per QALY gained. Assuming 90% persistence for romosozumab resulted in an ICER approximately mid-way between the ERG base-case and the company base-case at £267,533 per QALY gained. The scenario assuming romosozumab persistence per the ERG base-case and comparator persistence per the company base-case and the scenario assuming all treatments had persistence of 100% resulted in similar substantial increases in incremental QALYs and ICERs of approximately £40,000 per QALY gained (a decrease of approximately £443,000 in the ICER. Scenarios assuming persistence data based on trial data for all treatments and assuming romosozumab persistence equal to that of teriparatide resulted in negative incremental QALYs for romosozumab followed by alendronate, resulting in the treatment being dominated by alendronate.

Table 6.5: Scenario set 3 results: persistence

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Morley 2020 'All patients' for persistence with oral BPs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	81,333
90% persistence with romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	267,533
Romo persistence per ERG BC; Comparators per company BC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,315
Persistence based on trial data for all treatments	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Romo dominated
Romo persistence equal to teriparatide persistence	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Romo dominated

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
All treatments 100% persistence	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,539

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.4 Scenario set 4 results: patient support programme in place

The results of scenario analyses where it is assumed that the PSP is in place, are provided in Table 6.6. Assuming no administration costs for romosozumab had a minor impact on the results. In the scenario where the same assumption was combined with 90% persistence with romosozumab the ICER was almost halved.

Table 6.6: Scenario set 4 results: patient support programme in place

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
No admin. costs for romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	471,250
No admin. costs + 90% persistence with romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	260,533

Based on the ERG preferred version of the electronic model.[1]
admin. = administration; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.5 Scenario set 5 results: costs associated with fractures

The results of scenario analyses set 5 are provided in Table 6.7. In this scenario the total health care costs associated with fractures from Gutiérrez et al. 2011 and 2012 are applied, instead of the incremental costs of patients with fractures vs. those without from the same sources.[136, 137] The impact of this assumption on the model results was minimal.

Table 6.7: Scenario set 5 results: costs associated with fractures

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Scenario 5: total health care costs associated with fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	482,750

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.6 Scenario set 6 results: utility multipliers

The results of the utility multiplier scenarios are provided in Table 6.8. The ICER was very sensitive to the multipliers applied as the incremental QALYs in the ERG base-case are so small and, therefore, changes to incremental QALYs have a large impact on the ICER. Using the TA646 multipliers approximately doubled the incremental QALY gain to [redacted] from [redacted], which led to a substantial reduction in the ICER to £258,000 from £483,750. Conversely, using the multiplier from ID901led to a small decrease of approximately [redacted] in the incremental QALYs, but still increased the ICER by approximately £70,000 per QALY gained.

Table 6.8: Scenario set 6 results: utility multipliers

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Base-case multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
TA464 multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	258,000
ID901 multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	552,857

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.7 Scenario set 7 results: CV AEs

The results of the scenario in which CV AEs were removed from the ERG base-case are shown in Table 6.9. Removing the CV AEs led to a decrease in incremental costs and an increase in incremental QALYs, resulting in a decrease of approximately £173,000 in the ICER.

Table 6.9: Scenario set 7 results: CV AEs

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
CV AEs included (ERG)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
No CV AEs (company)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	310,917

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.8 Scenario set 8 results: treatment effect waning

Results for the treatment effect waning scenario are displayed in Table 6.10. The scenario in which 4 years of full treatment effect was assumed followed by a waning in effect for one more year resulted in a slight increase in incremental costs, and a slight reduction in incremental QALYs, which led to a substantial increase in the ICER of approximately £70,000 per QALY gained.

Table 6.10: Scenario set 8 results: treatment effect waning

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
No effect waning (BC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
4 years full effect then 1 year waning	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	554,714

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.9 Scenario set 9 results: excess mortality associated to fractures

Results for excess mortality scenarios are displayed in Table 6.11. The ERG base-case assumed excess mortality after hip fracture only. Including excess mortality also after vertebral fracture decreased the ICER by approximately £130,000 per QALY gained, due to an increase in incremental QALYs. The further addition of excess mortality due to NHNV had almost no impact on the ICER.

Table 6.11: Scenario set 9 results: excess mortality associated to fractures

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Hip only (ERG BC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Hip and vertebral	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	355,273
Hip, vertebral and NHNV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	354,545

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; NHNV = non-hip non-vertebral; QALY(s) = quality-adjusted life year(s)

6.3 ERG preferred assumptions

Tables 6.12 and 6.13 show the step-by-step changes made by the ERG to the company base-case alongside the cumulative and one-by-one impact of each change on the results, respectively. The change with the largest impact (by far) on the results was sourcing alendronate persistence estimates from Morley et al. 2020.[100] This highlights the importance of persistence parameters on the CE results. Other changes like including CV events in the analysis had a large impact on the cumulative base-case, because the ICER now was very sensitive given the small incremental QALYs, but not when this change is applied alone. The following three changes, when applied in isolation, resulted in an ICER that increased from below to above £20,000 per QALY gained: assuming 80% persistence with romosozumab (i.e., instead of 90%), assuming a daily cost of long-term care of £67 (i.e., instead of £112), and assuming incremental costs associated with fractures (i.e., of patients with fractures vs. those without, instead of total health care costs). The other changes, when applied in isolation, also resulted in increased ICERs but still remained below £20,000 per QALY gained.

Table 6.12: Incremental impact of ERG preferred assumptions (cumulative)

Preferred assumption (Section in ERG report)	Romosozumab 12 months / alendronate 48 months		Alendronate 48 months		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	262,429
+ Excess mortality only for hip fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,000
+ Daily LTC costs £67	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,000
+ Incremental fracture costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,750
+ CV events included	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	473,375
+ No PSP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	485,875
+ 2 GP visits per year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	484,250
+ UK general population mortality 2017 - 2019	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750

Based on the ERG preferred version of the electronic model.[1]
CV = cardiovascular; ERG = Evidence Review Group; GP = general practitioner; ICER = incremental cost effectiveness ratio; Incr. = incremental; LTC = long-term care; PSP = patient support programme; QALY(s) = quality-adjusted life year(s)

Table 6.13: Incremental impact of ERG preferred assumptions (one-by-one)

Preferred assumption (Section in ERG report)	Romosozumab 12 months / alendronate 48 months		Alendronate 48 months		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	162,391
+ Excess mortality only for hip fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,185
+ Daily LTC costs £67	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,476
+ Incremental fracture costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,398
+ CV events included	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,500
+ No PSP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,680
+ 2 GP visits per year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,117
+ UK general population mortality 2017 - 2019	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,903

Based on the ERG preferred version of the electronic model.[1]
CV = cardiovascular; ERG = Evidence Review Group; GP = general practitioner; ICER = incremental cost effectiveness ratio; Incr. = incremental; LTC = long-term care; PSP = patient support programme; QALY(s) = quality-adjusted life year(s).

6.4 Conclusions of the cost effectiveness section

The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported. Overall, the ERG does not have any major concerns regarding the searches but notes that no searches were conducted to identify health-state utility values (see Section 4.1.1 for more details), it is unclear whether empirical studies estimating utility values in this condition were missed as only included economic evaluations were searched for utility values.

The company developed a "de novo" Markov microsimulation model in Microsoft Excel. The model structure appears appropriate in general. However, the ERG's ability to step through and evaluate the model functionality was hindered by the fact that all model calculations are done in background VBA code that could not be changed. Therefore, the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. The CE analysis was performed in line with the NICE Reference case in terms of perspective, time horizon and discounting.[77]

The population in the Final Scope by NICE is defined as "Postmenopausal women with severe osteoporosis at high risk of fracture", in line with romosozumab marketing authorisation. The modelled population in the CS is assumed to consist of patients who are at imminent risk of another fragility fracture i.e. have had a MOF within the prior 24 months. An important difference between the ARCH ITT population and the modelled population is that ARCH included patients who previously sustained a fracture regardless of recency, whereas for the modelled population it is assumed that a previous fracture was sustained within 24 months prior to the start of treatment. In the ARCH ITT population, [redacted] of patients suffered a MOF within 24 months prior to randomisation. The differences between the definition in the NICE final scope, the ITT population from ARCH that was used to inform treatment effectiveness inputs for the company's base-case analysis, and the definition of the modelled population in the CS, present a key issue of uncertainty. It is not clear whether the term 'high risk' as used in the definitions in the NICE final scope and EMA marketing authorisation corresponds to the same definition that is used in the literature for the categorisation of fracture risk to guide choice of treatment. It is, therefore, uncertain whether the ITT population results are representative for the population in the CS and whether these are generalisable to the target population of romosozumab.

The modelled intervention consisted of a 12-month course of romosozumab, followed by a 48-month course of alendronate. The comparators that were used in the company base-case consist of a 60-month course of alendronate and no treatment. Additionally, the company included additional comparators (teriparatide, denosumab, risedronate, zoledronate, and raloxifene) as scenario analyses. All treatments considered by the company were listed in the NICE scope, except for ibandronic acid, for which the company identified no trials at its licensed dose and, therefore, it could not be included in the analyses. Given the uncertainty regarding the relevant population for this appraisal in Section 2.1, as previously described, and the lack of clarity of current guidelines, there is also uncertainty regarding the appropriateness and relevance of the included comparators. In particular, if high risk is differentiated from very high risk, then alendronate might be the most appropriate comparator, but if high risk includes very high risk, then other comparators might be appropriate.

In the model, the risk of fractures in patients with severe osteoporosis who had a MOF in the prior 24 months is estimated using three components: general population risk of fractures, increased risk of fractures associated with osteoporosis, imminent risk of subsequent fractures following an index fracture. The general population risk of hip fractures was sourced from Singer et al. 1998 and the same source was used to estimate the incidence of vertebral fractures using the ratio of hip to vertebral fractures from a Swedish study.[81, 83] To estimate the incidence of NHNV fractures, Singer et al. 1998 was used for forearm fractures and the same approach that was used to estimate the incidence of vertebral fractures was applied to the other types of fractures that are included in NHNV fractures.[81] No changes were applied by the ERG, but the ERG did note some uncertainty regarding the validity of estimates of fracture incidence that was related to the stability over time of fracture incidence and the assumption that ratios between different types of fractures as found in Sweden also apply to the UK. The increased risk of fracture due to osteoporosis, relative to the general population, was estimated using FRAX whilst excluding prior fracture as a clinical risk factor. Finally, the additional risk of experiencing a subsequent fracture after an index fracture was based on the maximum of the 'imminent

risk', sourced from Söreskog et al. 2020,[85] or the additional risk from FRAX whilst including prior fracture as a risk factor.

Efficacy estimates for romosozumab vs. alendronate were based on ARCH data, by reconstructing patient-level data from published Kaplan-Meier curves and then fitting parametric distributions in order to calculate time-dependent hazard rates. These (survival data) analyses were not presented by the company. While the methods used seem appropriate, the ERG cannot assess whether the distributions were properly fitted and cannot explore the impact of using alternative distributions on the model results. In analyses vs. other comparators, efficacy was estimated using an NMA in which treatment effects were estimated on the trial ITT population. Limitations of the NMA were discussed in the clinical effectiveness sections of the report (e.g., Section 3.6).

The company modelled persistence with osteoporosis therapies based on the assumption that real-world persistence with romosozumab would equal persistence as found in ARCH and that persistence with alendronate following romosozumab would be 85% of real-world persistence with denosumab from Li et al. 2012. Persistence with alendronate alone was also based on Li et al. 2012.[96] The ERG identified Morley et al. 2020 [100] as a more recent source of persistence estimates that is effectively an update of the study by Li et al. 2012 (both based on GPRD), that they preferred to use for their base-case analysis to inform persistence with alendronate, using estimates from non-naïve patients for alendronate after romosozumab and estimates from naïve patients for alendronate alone. This change, when applied in isolation of the other ERG changes, resulted in the largest impact on the CE results and increased the ICER by nearly ten-fold.

The company assumed that anti-fracture efficacy persists for a period of time (offset time) after treatment is discontinued in patients with osteoporosis.[102] A dynamic offset time equal to time on treatment is assumed for the base-case. During the offset time the fracture risk reduction is assumed to decline linearly to zero. The efficacy of the last treatment given to the patient in the sequence was used for the offset time. This approach is recommended by the ESCEO and IOF guidelines, and has been used in other published health economic studies and romosozumab HTA submissions to the SMC and TLV.[78, 103, 104] Therefore, the ERG considers the company's approach appropriate. Scenarios with fixed offset time can be deemed as exploratory. As described in Key issue 2, scenarios with shorter duration of the dynamic offset of the treatment effect could be of interest of being explored. However, the ERG was unable to run this type of scenarios, which are expected to increase the ICER. Finally, the ERG would like to note that residual effects for zoledronate could be longer than those assumed by the company.[112] The ERG was unable to change the model to incorporate this assumption. Cost effectiveness results including zoledronate as comparator might be underestimating the ICER (even though in these scenarios zoledronate was dominant over romosozumab).

Mortality is captured in the model in three ways: age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor.[1] It is unclear why the company used UK Life Tables from 2012 to 2014. In the ERG base-case, the most recent version (2017 to 2019) was used.[115] When a patient sustains a fracture, the relative risk of death compared with the non-fractured population is applied to the normal population risk, and the relative risk was down-adjusted to 30% to adjust for higher frailty (i.e., increased risk of death due to other reasons than the fracture itself) in the fractured population.[1, 94, 114] The company included in the base-case mortality related to hip, clinical vertebral, and NHNV fractures. Following the expert reviewers comments to the ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] the ERG prefers to include excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNV fractures were explored by the ERG.

The company only included GI AEs associated with bisphosphonates in their base-case. An imbalance in serious adjudicated CV AEs was observed in the ARCH trial, which led to romosozumab being contraindicated for patients with previous MI or stroke. The company chose to exclude CV AEs from their base-case due to this contraindication. However, the ERG considered that those CV events which occurred in patients without a history of MI or stroke should be included as they would not be avoided by the contraindication. These CV events in patients without history were therefore included in the ERG base-case.

Utilities for fracture health states within the model were estimated using fracture multipliers from the international ICUROS study, multiplied with UK age adjusted general population utility values. Separate multipliers were provided for hip, vertebral and NHNV fractures during the first (acute) and subsequent (chronic) years after fracture. This utility approach follows previous appraisals in osteoporosis, although some differences in multipliers were observed across appraisals. Multipliers from other available NICE appraisals were used in scenarios to examine the impact of differences on results. The ERG was unsure about the appropriateness of several assumptions in the utility analysis. In TA464, only one acute multiplier could be applied at any one time, while in this model two acute multipliers could be applied at once. Additionally, the ERG was unclear whether the assumption that chronic fracture multipliers were used for the remainder patients' lifetimes was supported by evidence. The company presented some evidence up to 5 years post fracture, but none beyond. However, the ERG was unable to test the impact of these assumptions, given that they could not access the VBA code in the validated version of the model on which analyses had to be conducted.

The following cost categories were included in the analysis: drug acquisition costs, drug administration costs, disease management costs, costs associated with fractures (i.e., hip fractures, vertebral fractures, and NHNV fractures), long-term care costs after a hip fracture, and costs for the treatment of GIAEs. The drug acquisition costs for romosozumab are £427.75 per set of two pre-filled disposable 1.17 ml injections of 90 mg/ml at list price or [redacted] including a PAS discount, resulting in an annual cost of £5,133 at list price, or [redacted] including the PAS discount. No drug administration costs were included for romosozumab, which the company justified by referring to their plans to set up a PSP that includes homecare service, an adherence support program, and training of injection techniques. Drug administration costs were included in the model only for patients receiving denosumab or zoledronate. However, since the PSP is not yet in place, the ERG preferred to include the costs for administration of romosozumab. For disease management costs, the ERG preferred the assumption that monitoring of osteoporosis therapies requires physician visits once a year to twice a year, in line with Hilligsmann et al. 2019.[78] The ERG preferred to use estimates of costs associated with fractures that were based on incremental costs of patients with fractures vs. patients without, rather than total costs of patients with fractures that were used by the company, in line with NICE TA464 and ID901.[11, 87] Lastly, the ERG preferred to use a different estimate of long-term costs based on the estimate as used in TA464.[11]

The company's deterministic base-case results indicate that romosozumab followed by alendronate is more costly and more effective than alendronate alone, with incremental QALYs of [redacted] and incremental costs of [redacted], resulting in an ICER of £16,660 per QALY gained. In the fully incremental analysis, no treatment was dominated by alendronate alone. The company's PSA results were more or less in line with their deterministic results, with a probabilistic ICER of £14,537 per QALY gained. At a threshold of £30,000 per QALY gained, the estimated probability that romosozumab is a cost-effective alternative to alendronate alone or no treatment is [redacted] and [redacted] respectively. The company's DSA shows that model results are sensitive to varying the time horizon, persistence, start age, changes in treatment effect of romosozumab/alendronate and alendronate alone, and utility multipliers for hip,

vertebral and NHNV fracture. Company scenario analyses highlighted the sensitivity of results to persistence assumptions and the removal of imminent risk in the calculation of fracture incidence.

The ERG base-case differed from the company base-case in a number of elements including: the assumed persistence rates for romosozumab and alendronate; assumed excess mortality after vertebral and NHNV fractures; incremental fracture and daily LTC costs; inclusion of CV AEs and PSP costs; number of GP visits per year and the source of UK general population mortality rates. The ERG change that had by far the most impact on results when applied in isolation was assuming the persistence estimate for alendronate from Morley et al. 2020. Which increased the company base-case ICER ten-fold, from £16,660 to £162,391 per QALY gained. The next most influential parameters reducing the daily LTC cost, assuming 80% persistence on romosozumab and reducing the incremental fracture costs, all of which when applied individually took the ICER over £20,000 per QALY gained.

The ERG deterministic base-case resulted in higher incremental costs ([redacted] vs. [redacted]) and substantially lower incremental QALYs ([redacted] vs. [redacted]) which resulted in a high ICER of £483,750. A PSA on the ERG base-case could not be run as the model continued to crash. However, given the size of the ICER, it is likely that at the usual threshold range of £20,000 to £30,000 per QALY gained, the probability of romosozumab being cost effective would be [redacted]%. Scenario analyses run on the ERG preferred assumptions showed that model results were most sensitive to assumed rates of persistence, however, scenarios surrounding utility multipliers, treatment effect waning, excess mortality due to fractures and inclusion of CV AEs and PSP also had large impacts on the ICER, which was very sensitive to changes in the small incremental QALYs. When various alternative comparators were included in the analysis, romosozumab was dominated by zoledronate. In this situation, the only relevant comparison was zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab, showed that all ICERs were above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 18 months, teriparatide 24 months and teriparatide followed by alendronate, which are dominated by romosozumab.

Regarding validation, the model adheres in general to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] Since 2017, the model has been involved in several iterations of quality assessment including the NICE PRIMA. In line with this assessment, the ERG considers that review would be better facilitated if calculations were performed in the model work sheets, instead of being hard coded in VBA. Some discrepancies between the model results and the trackers summary were found, which could not all be traced to their source in the VBA code. An error was found with regards to the presence of previous fractures, and it is not clear is this has any impact on the results. Additionally, the model seems to be extremely demanding regarding the computational power needed to run within reasonable time. Even a deterministic run would take more than 20 minutes. This makes the validation process extra difficult and for this reason, the ERG was not able to validate the results of some of the scenarios presented by the company and did not succeed in running any PSA. The main concerns of the ERG relate to the validity of the baseline fracture incidences as noted above. Also, validation was presented against Swedish data only because for example UK data on fractures and risk factors such as BMD were not available. Therefore, it is uncertain whether the validity results can be generalised to the UK. Finally, comparisons with other technology appraisals were not presented. Therefore, it is not possible to quantify whether the results in the CS are in line with those in previous appraisals.

The same issues identified in the clinical effectiveness section are carried over in the economic analyses. The model results are affected by the limitations of the NMAs, and they should be interpreted in a

similar way as the results of the NMAs: with caution. If additional data are identified to reduce bias in the NMAs, this would also reduce the uncertainty around the model results. However, it is uncertain what the effect on the CE estimates might be.

In conclusion, in contrast to the company's base-case that resulted in an ICER of £16,660 per QALY gained, the ERG preferred base-case results in an ICER of £483,750 per QALY. This difference is mainly caused by different assumptions regarding the persistence with alendronate. The high value of the ICER can further be explained by the higher incremental costs ([redacted] vs. [redacted]) and substantially lower incremental QALYs ([redacted] vs. [redacted]) of the ERG preferred base-case vs. the company's base-case, respectively.

7 REFERENCES

[1] UCB Pharma Ltd. *Romosozumab for treating postmenopausal severe osteoporosis [ID3936]. Document B. Company evidence submission for romosozumab. National Institute of Health and Care Excellence. Single technology appraisal (STA)*: UCB Pharma Ltd, 2021

[2] National Institute for Health and Care Excellence. *Romosozumab for treating severe osteoporosis. Health Technology Appraisal. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/gid-ta10828/documents/final-scope

[3] Saag KG, Petersen J, Brandi ML, Karaplis AC, Lorentzon M, Thomas T, et al. Romosozumab or alendronate for fracture prevention in women with osteoporosis. *N Engl J Med* 2017;377(15):1417-27.

[4] UCB Pharma Limited. EVENITY 105 mg solution for injection in pre-filled pen (romosozumab). Summary of product characteristics [Internet]. electronic Medicines Compendium (eMC), 2021 [accessed 1.7.21]. Available from: https://www.medicines.org.uk/emc/product/10956/smpc

[5] Amgen Inc. *STRUCTURE. Clinical study report: 20080289. An open-label, randomized, teriparatide-controlled study to evaluate the effect of treatment with romosozumab in postmenopausal women with osteoporosis previously treated with bisphosphonate therapy. UCB Data on File*, 2015

[6] Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761-7.

[7] Alswat KA. Gender disparities in osteoporosis. *J Clin Med Res* 2017;9(5):382-7.

[8] UCB Pharma Ltd. *Romosozumab for treating postmenopausal severe osteoporosis [ID3936]. Appendices. Company evidence submission for romosozumab. National Institute of Health and Care Excellence. Single technology appraisal (STA)*: UCB Pharma Ltd, 2021

[9] UCB Pharma Ltd. *Romosozumab for treating severe osteoporosis [ID3936]. Clarification questions. National Institute of Health and Care Excellence. Single technology appraisal (STA). Company response to clarification*: UCB Pharma Ltd, 2021

[10] Golder S, Peryer G, Loke YK. Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *J Clin Epidemiol* 2019;113:36-43.

[11] National Institute for Health and Care Excellence. *Bisphosphonates for treating osteoporosis. NICE technology appraisal guidance 464 [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta464

[12] Bai H, Jing D, Guo A, Yin S. Randomized controlled trial of zoledronic acid for treatment of osteoporosis in women. *J Int Med Res* 2013;41(3):697-704.

[13] Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med* 2007;356(18):1809-22.

[14] Morii H, Ohashi Y, Taketani Y, Fukunaga M, Nakamura T, Itabashi A, et al. Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. *Osteoporos Int* 2003;14(10):793-800.

[15] Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med* 2009;361(8):756-65.

[16] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021) [Internet]*: Cochrane, 2021 [accessed 1.7.21] Available from: https://training.cochrane.org/handbook

[17] Cosman F, Crittenden DB, Adachi JD, Binkley N, Czerwinski E, Ferrari S, et al. Romosozumab treatment in postmenopausal women with osteoporosis. *N Engl J Med* 2016;375(16):1532-43.

[18] Langdahl BL, Libanati C, Crittenden DB, Bolognese MA, Brown JP, Daizadeh NS, et al. Romosozumab (sclerostin monoclonal antibody) versus teriparatide in postmenopausal women with osteoporosis transitioning from oral bisphosphonate therapy: a randomised, open-label, phase 3 trial. *Lancet* 2017;390(10102):1585-94.

[19] Amgen Inc. *BRIDGE. Clinical study report: 20110174. A multicenter, randomized, double-blind, placebo-controlled study to compare the efficacy and safety of romosozumab with placebo in men with osteoporosis. UCB Data on File*, 2016

[20] Amgen Inc. *ARCH. Clinical study report: 20110142. A multicenter, international, randomized, double-blind, alendronate-controlled study to determine the efficacy and safety of romosozumab in the treatment of postmenopausal women with osteoporosis. UCB Data on File*, 2017

[21] Amgen. Study to determine the efficacy and safety of romosozumab in the treatment of postmenopausal women with osteoporosis (ARCH). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2012-2017 [cited 2021 Aug 04]. Available from: https://clinicaltrials.gov/ct2/show/NCT01631214. NLM Identifier: NCT01631214

[22] Amgen. Efficacy and safety of romosozumab treatment in postmenopausal women with osteoporosis (FRAME). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2012-2016 [cited 2021 Jul 01]. Available from: https://clinicaltrials.gov/ct2/show/NCT01575834. NLM Identifier: NCT01575834

[23] Miller PD, Hattersley G, Riis BJ, Williams GC, Lau E, Russo LA, et al. Effect of abaloparatide vs placebo on new vertebral fractures in postmenopausal women with osteoporosis: a randomized clinical trial. *JAMA* 2016;316(7):722-33.

[24] Chao M, Hua Q, Yingfeng Z, Guang W, Shufeng S, Yuzhen D, et al. Study on the role of zoledronic acid in treatment of postmenopausal osteoporosis women. *Pak J Med Sci* 2013;29(6):1381-4.

[25] Dursun N, Dursun E, Yalçin S. Comparison of alendronate, calcitonin and calcium treatments in postmenopausal osteoporosis. *Int J Clin Pract* 2001;55(8):505-9.

[26] Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA* 1998;280(24):2077-82.

[27] Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. *Lancet* 1996;348(9041):1535-41.

[28] Pols HA, Felsenberg D, Hanley DA, Stepán J, Muñoz-Torres M, Wilkin TJ, et al. Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in

postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. *Osteoporos Int* 1999;9(5):461-8.

[29] Hadji P, Zanchetta JR, Russo L, Recknor CP, Saag KG, McKiernan FE, et al. The effect of teriparatide compared with risedronate on reduction of back pain in postmenopausal women with osteoporotic vertebral fractures. *Osteoporos Int* 2012;23(8):2141-50.

[30] Liberman UA, Weiss SR, Bröll J, Minne HW, Quan H, Bell NH, et al. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. *N Engl J Med* 1995;333(22):1437-43.

[31] Liu JL, Zhu HM, Huang QR, Zhang ZL, Li HL, Qin YJ, et al. Effects of raloxifene hydrochloride on bone mineral density, bone metabolism and serum lipids in Chinese postmenopausal women with osteoporosis: a multi-center, randomized, placebo-controlled clinical trial. *Chinese Medical Journal* 2004;117(7):1029-35.

[32] Lufkin EG, Whitaker MD, Nickelsen T, Argueta R, Caplan RH, Knickerbocker RK, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. *J Bone Miner Res* 1998;13(11):1747-54.

[33] Delmas PD, Ensrud KE, Adachi JD, Harper KD, Sarkar S, Gennari C, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. *J Clin Endocrinol Metab* 2002;87(8):3609-17.

[34] Neer RM, Arnaud CD, Zanchetta JR, Prince R, Gaich GA, Reginster JY, et al. Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N Engl J Med* 2001;344(19):1434-41.

[35] Hadji P, Gamerdinger D, Spieler W, Kann PH, Loeffler H, Articus K, et al. Rapid Onset and Sustained Efficacy (ROSE) study: results of a randomised, multicentre trial comparing the effect of zoledronic acid or alendronate on bone metabolism in postmenopausal women with low bone mass. *Osteoporos Int* 2012;23(2):625-33.

[36] Ensrud KE, Stock JL, Barrett-Connor E, Grady D, Mosca L, Khaw KT, et al. Effects of raloxifene on fracture risk in postmenopausal women: the Raloxifene Use for the Heart Trial. *J Bone Miner Res* 2008;23(1):112-20.

[37] Silverman SL, Christiansen C, Genant HK, Vukicevic S, Zanchetta JR, de Villiers TJ, et al. Efficacy of bazedoxifene in reducing new vertebral fracture risk in postmenopausal women with osteoporosis: results from a 3-year, randomized, placebo-, and active-controlled clinical trial. *J Bone Miner Res* 2008;23(12):1923-34.

[38] Kendler DL, Marin F, Zerbini CAF, Russo LA, Greenspan SL, Zikan V, et al. Effects of teriparatide and risedronate on new fractures in post-menopausal women with severe osteoporosis (VERO): a multicentre, double-blind, double-dummy, randomised controlled trial. *Lancet* 2018;391(10117):230-40.

[39] Reginster J, Minne HW, Sorensen OH, Hooper M, Roux C, Brandi ML, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. *Osteoporos Int* 2000;11(1):83-91.

[40] Harris ST, Watts NB, Genant HK, McKeever CD, Hangartner T, Keller M, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal

osteoporosis: a randomized controlled trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. *JAMA* 1999;282(14):1344-52.

[41] Nakamura T, Fukunaga M, Nakano T, Kishimoto H, Ito M, Hagino H, et al. Efficacy and safety of once-yearly zoledronic acid in Japanese patients with primary osteoporosis: two-year results from a randomized placebo-controlled double-blind study (ZOledroNate treatment in Efficacy to osteoporosis; ZONE study). *Osteoporos Int* 2017;28(1):389-98.

[42] Adami S, Passeri M, Ortolani S, Broggini M, Carratelli L, Caruso I, et al. Effects of oral alendronate and intranasal salmon calcitonin on bone mass and biochemical markers of bone turnover in postmenopausal women with osteoporosis. *Bone* 1995;17(4):383-90.

[43] Adami S, San Martin J, Munoz-Torres M, Econs MJ, Xie L, Dalsky GP, et al. Effect of raloxifene after recombinant teriparatide [hPTH(1-34)] treatment in postmenopausal women with osteoporosis. *Osteoporos Int* 2008;19(1):87-94.

[44] Aki S, Gunay Gulbabab R, Eskiyurta N. Effect of alendronate on bone density and bone markers in postmenopausal osteoporosis. *J Back Musculoskelet Rehabil* 2004;17(2003/2004):27-31.

[45] McClung MR, Lewiecki EM, Cohen SB, Bolognese MA, Woodson GC, Moffett AH, et al. Denosumab in postmenopausal women with low bone mineral density. *N Engl J Med* 2006;354(8):821-31.

[46] Tsai JN, Uihlein AV, Lee H, Kumbhani R, Siwila-Sackman E, McKay EA, et al. Teriparatide and denosumab, alone or combined, in women with postmenopausal osteoporosis: the DATA study randomised trial. *Lancet* 2013;382(9886):50-6.

[47] Brown JP, Prince RL, Deal C, Recker RR, Kiel DP, de Gregorio LH, et al. Comparison of the effect of denosumab and alendronate on BMD and biochemical markers of bone turnover in postmenopausal women with low bone mass: a randomized, blinded, phase 3 trial. *J Bone Miner Res* 2009;24(1):153-61.

[48] Bone HG, Bolognese MA, Yuen CK, Kendler DL, Wang H, Liu Y, et al. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. *J Clin Endocrinol Metab* 2008;93(6):2149-57.

[49] Luckey M, Kagan R, Greenspan S, Bone H, Kiel RD, Simon J, et al. Once-weekly alendronate 70 mg and raloxifene 60 mg daily in the treatment of postmenopausal osteoporosis. *Menopause* 2004;11(4):405-15.

[50] Sambrook PN, Geusens P, Ribot C, Solimano JA, Ferrer-Barriendos J, Gaines K, et al. Alendronate produces greater effects than raloxifene on bone density and bone turnover in postmenopausal women with low bone density: results of EFFECT (Efficacy of FOSAMAX versus EVISTA Comparison Trial) International. *J Intern Med* 2004;255(4):503-11.

[51] Eastell R, Nickelsen T, Marin F, Barker C, Hadji P, Farrerons J, et al. Sequential treatment of severe postmenopausal osteoporosis after teriparatide: final results of the randomized, controlled European Study of Forsteo (EUROFORS). *J Bone Miner Res* 2009;24(4):726-36.

[52] Rosen CJ, Hochberg MC, Bonnick SL, McClung M, Miller P, Broy S, et al. Treatment with once-weekly alendronate 70 mg compared with once-weekly risedronate 35 mg in women with postmenopausal osteoporosis: a randomized double-blind study. *J Bone Miner Res* 2005;20(1):141-51.

[53] Reid DM, Hosking D, Kendler D, Brandi ML, Wark JD, Marques-Neto JF, et al. A comparison of the effect of alendronate and risedronate on bone mineral density in postmenopausal women with osteoporosis: 24-month results from FACTS-International. *Int J Clin Pract* 2008;62(4):575-84.

[54] Fogelman I, Ribot C, Smith R, Ethgen D, Sod E, Reginster JY. Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab* 2000;85(5):1895-900.

[55] Grey A, Bolland M, Wattie D, Horne A, Gamble G, Reid IR. Prolonged antiresorptive activity of zoledronate: a randomized, controlled trial. *J Bone Miner Res* 2010;25(10):2251-5.

[56] Johnell O, Scheele WH, Lu Y, Reginster JY, Need AG, Seeman E. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. *J Clin Endocrinol Metab* 2002;87(3):985-92.

[57] McClung MR, Bolognese MA, Sedarati F, Recker RR, Miller PD. Efficacy and safety of monthly oral ibandronate in the prevention of postmenopausal bone loss. *Bone* 2009;44(3):418-22.

[58] McClung MR, Grauer A, Boonen S, Bolognese MA, Brown JP, Diez-Perez A, et al. Romosozumab in postmenopausal women with low bone mineral density. *N Engl J Med* 2014;370(5):412-20.

[59] Amgen Inc. A study to evaluate the safety and efficacy of denosumab compared with zoledronic acid in postmenopausal women with osteoporosis previously treated with oral bisphosphonates. EUCTR2012-001821-28-ES. In: WHO International Clinical Trials Registry Platform (ICTRP) [Internet]. Geneva: World Health Organization (WHO), 2012 [accessed 4.8.21].

[60] Miller PD, Epstein S, Sedarati F, Reginster JY. Once-monthly oral ibandronate compared with weekly oral alendronate in postmenopausal osteoporosis: results from the head-to-head MOTION study. *Curr Med Res Opin* 2008;24(1):207-13.

[61] Valimaki MJ, Farrerons-Minguella J, Halse J, Kroger H, Maroni M, Mulder H, et al. Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. *Clin Ther* 2007;29(9):1937-49.

[62] Bonnick S, Broy S, Kaiser F, Teutsch C, Rosenberg E, DeLucca P, et al. Treatment with alendronate plus calcium, alendronate alone, or calcium alone for postmenopausal low bone mineral density. *Curr Med Res Opin* 2007;23(6):1341-9.

[63] Eastell R, Nagase S, Ohyama M, Small M, Sawyer J, Boonen S, et al. Safety and efficacy of the cathepsin K inhibitor ONO-5334 in postmenopausal osteoporosis: the OCEAN study. *J Bone Miner Res* 2011;26(6):1303-12.

[64] Recknor C, Czerwinski E, Bone HG, Bonnick SL, Binkley N, Palacios S, et al. Denosumab compared with ibandronate in postmenopausal women previously treated with bisphosphonate therapy: a randomized open-label trial. *Obstet Gynecol* 2013;121(6):1291-9.

[65] Reid IR, Bolland M, Wong S, et al. A randomized controlled trial of annual low dose zoledronate in postmenopausal women. *Osteoporos Int* 2011;2:S526.

[66] Roux C, Hofbauer LC, Ho PR, Wark JD, Zillikens MC, Fahrleitner-Pammer A, et al. Denosumab compared with risedronate in postmenopausal women suboptimally adherent to alendronate therapy: efficacy and safety results from a randomized open-label study. *Bone* 2014;58:48-54.

[67] Bock O, Borst H, Beller G, Armbrecht G, Degner C, Martus P, et al. Impact of oral ibandronate 150 mg once monthly on bone structure and density in post-menopausal osteoporosis or osteopenia derived from in vivo muCT. *Bone* 2012;50(1):317-24.

[68] Kendler DL, Roux C, Benhamou CL, Brown JP, Lillestol M, Siddhanti S, et al. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women transitioning from alendronate therapy. *J Bone Miner Res* 2010;25(1):72-81.

[69] Tan W, Sun J, Zhou L, Li Y, Wu X. Randomized trial comparing efficacies of zoledronate and alendronate for improving bone mineral density and inhibiting bone remodelling in women with post-menopausal osteoporosis. *J Clin Pharm Ther* 2016;41(5):519-23.

[70] Tucci JR, Tonino RP, Emkey RD, Peverly CA, Kher U, Santora AC, 2nd. Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am J Med* 1996;101(5):488-501.

[71] Um MJ, Cho EA, Jung H. Combination therapy of raloxifene and alendronate for treatment of osteoporosis in elderly women. *J Menopausal Med* 2017;23(1):56-62.

[72] Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. *Syst Rev* 2021;10(1):39.

[73] Söreskog E, Lindberg I, Kanis JA, Åkesson KE, Willems D, Lorentzon M, et al. Cost-effectiveness of romosozumab for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture in Sweden. *Osteoporos Int* 2021;32(3):585-94.

[74] Soreskog E, Borgstrom F, Lindberg I, Strom O, Willems D, Libanati C, et al. A novel economic framework to assess the cost-effectiveness of bone-forming agents in the prevention of fractures in patients with osteoporosis. *Osteoporos Int* 2021;32(7):1301-1311.

[75] Quantify. *Quantify Research Data on File. Swedish Osteoporosis Research Database*, 2017

[76] National Osteoporosis Guidelines Group. *NOGG 2017: Clinical guideline for the prevention and treatment of osteoporosis [Internet]*, 2017 [accessed 1.7.21] Available from: https://www.sheffield.ac.uk/NOGG/NOGG%20Guideline%202017.pdf

[77] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 (2018 revision) [Internet]*. London: NICE, 2013 (revised 2018 [accessed 1.7.21] Available from: https://www.nice.org.uk/process/pmg9/

[78] Hiligsmann M, Reginster JY, Tosteson ANA, Bukata SV, Saag KG, Gold DT, et al. Recommendations for the conduct of economic evaluations in osteoporosis: outcomes of an experts' consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the US branch of the International Osteoporosis Foundation. *Osteoporos Int* 2019;30(1):45-57.

[79] Svedbom A, Hernlund E, Ivergard M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: a compendium of country-specific reports. *Arch Osteoporos* 2013;8:137.

[80] Hernlund E, Svedbom A, Ivergard M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch Osteoporos* 2013;8:136.

[81] Singer BR, McLauchlan GJ, Robinson CM, Christie J. Epidemiology of fractures in 15,000 adults: the influence of age and gender. *J Bone Joint Surg Br* 1998;80(2):243-8.

[82] van der Velde RY, Wyers CE, Curtis EM, Geusens P, van den Bergh JPW, de Vries F, et al. Secular trends in fracture incidence in the UK between 1990 and 2012. *Osteoporos Int* 2016;27(11):3197-3206.

[83] Kanis JA, Johnell O, Oden A, Sembo I, Redlund-Johnell I, Dawson A, et al. Long-term risk of osteoporotic fracture in Malmö. *Osteoporos Int* 2000;11(8):669-74.

[84] Kanis JA, Johnell O, Oden A, De Laet C, Oglesby A, Jönsson B. Intervention thresholds for osteoporosis. *Bone* 2002;31(1):26-31.

[85] Soreskog E, Strom O, Spangeus A, Akesson KE, Borgstrom F, Banefelt J, et al. Risk of major osteoporotic fracture after first, second and third fracture in Swedish women aged 50years and older. *Bone* 2020;134:115286.

[86] Banefelt J, Akesson KE, Spangeus A, Ljunggren O, Karlsson L, Strom O, et al. Risk of imminent fracture following a previous fracture in a Swedish database study. *Osteoporos Int* 2019;30(3):601-9.

[87] National Institute for Health and Care Excellence. *Assessment Group Report: Non-bisphosphonates for treating osteoporosis [ID901]*. Sheffield: School of Health and Related Research (ScHARR), University of Sheffield, 2019

[88] Simpson EL, Martyn-St James M, Hamilton J, Wong R, Gittoes N, Selby P, et al. Clinical effectiveness of denosumab, raloxifene, romosozumab, and teriparatide for the prevention of osteoporotic fragility fractures: a systematic review and network meta-analysis. *Bone* 2020;130:115081.

[89] Willems D, Javaid MK, Pinedo-Villaneuva R, Libanati C, Yehoshua A, Charokopou M, et al. UCB Data on File - SLR NMA. The importance of time-point specific indirect treatment comparisons of osteoporosis treatments: a systematic literature review (SLR) and network meta-analysis (NMA). *Clin Ther* 2021.

[90] Kanis JA, Oden A, Johnell O, Jonsson B, de Laet C, Dawson A. The burden of osteoporotic fractures: a method for setting intervention thresholds. *Osteoporos Int* 2001;12(5):417-27.

[91] Li L, Roddam A, Gitlin M, Taylor A, Shepherd S, Jick S. Concurrent Oral 8 – Innovations [OP54–OP58]: OP54. Non-persistence to anti-osteoporosis medications in the UK using the General Practice Research Database (GPRD). *Rheumatology (Oxford)* 2010;49(Suppl 1):i23-i25.

[92] Solomon DH, Avorn J, Katz JN, Finkelstein JS, Arnold M, Polinski JM, et al. Compliance with osteoporosis medications. *Arch Intern Med* 2005;165(20):2414-9.

[93] Huybrechts KF, Ishak KJ, Caro JJ. Assessment of compliance with osteoporosis treatment and its consequences in a managed care population. *Bone* 2006;38(6):922-8.

[94] Borgström F, Johnell O, Kanis JA, Oden A, Sykes D, Jönsson B. Cost effectiveness of raloxifene in the treatment of osteoporosis in Sweden: an economic evaluation based on the MORE study. *Pharmacoeconomics* 2004;22(17):1153-65.

[95] Davis S, Martyn-St James M, Sanderson J, Stevens J, Goka E, Rawdin A, et al. *Bisphosphonates for preventing osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161). Technology assessment report: final report to the National Institute for Health and Care Excellence [Internet]*. London: National Institute for Health and Care Excellence, 2015

[accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta464/documents/osteoporosis-prevention-bisphosphonates-inc-part-rev-ta160-ta161-id782-assessment-report2

[96] Li L, Roddam A, Gitlin M, Taylor A, Shepherd S, Shearer A, et al. Persistence with osteoporosis medications among postmenopausal women in the UK General Practice Research Database. *Menopause* 2012;19(1):33-40.

[97] Karlsson L, Lundkvist J, Psachoulia E, Intorcia M, Strom O. Persistence with denosumab and persistence with oral bisphosphonates for the treatment of postmenopausal osteoporosis: a retrospective, observational study, and a meta-analysis. *Osteoporos Int* 2015;26(10):2401-11.

[98] McHorney CA, Spain CV. Frequency of and reasons for medication non-fulfillment and non-persistence among American adults with chronic disease in 2008. *Health Expect* 2011;14(3):307-20.

[99] National Institute for Health and Care Excellence. *Osteoporosis: assessing the risk of fragility fracture. NICE clinical guideline 146 [Internet]*. London: National Institute for Health and Care Excellence, 2012 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/cg146

[100] Morley J, Moayyeri A, Ali L, Taylor A, Feudjo-Tepie M, Hamilton L, et al. Persistence and compliance with osteoporosis therapies among postmenopausal women in the UK Clinical Practice Research Datalink. *Osteoporos Int* 2020;31(3):533-45.

[101] Vanelli M, Pedan A, Liu N, Hoar J, Messier D, Kiarsis K. The role of patient inexperience in medication discontinuation: a retrospective analysis of medication nonpersistence in seven chronic illnesses. *Clin Ther* 2009;31(11):2628-52.

[102] Kanis JA, McCloskey EV, Jonsson B, Cooper A, Ström O, Borgström F. An evaluation of the NICE guidance for the prevention of osteoporotic fragility fractures in postmenopausal women. *Arch Osteoporos* 2010;5(1-2):19-48.

[103] Jonsson B, Strom O, Eisman JA, Papaioannou A, Siris ES, Tosteson A, et al. Cost-effectiveness of Denosumab for the treatment of postmenopausal osteoporosis. *Osteoporos Int* 2011;22(3):967-82.

[104] Tandvårds- och läkemedelsförmånsverket (The Dental and Pharmaceutical Benefits Agency). *Underlag för beslut om subvention - Nyansökan [Decision for reimbursement]*, 2020

[105] Black DM, Schwartz AV, Ensrud KE, Cauley JA, Levis S, Quandt SA, et al. Effects of continuing or stopping alendronate after 5 years of treatment: the Fracture Intervention Trial Long-term Extension (FLEX): a randomized trial. *JAMA* 2006;296(24):2927-38.

[106] Black DM, Reid IR, Boonen S, Bucci-Rechtweg C, Cauley JA, Cosman F, et al. The effect of 3 versus 6 years of zoledronic acid treatment of osteoporosis: a randomized extension to the HORIZON-Pivotal Fracture Trial (PFT). *J Bone Miner Res* 2012;27(12):243-54.

[107] Lindsay R, Scheele WH, Neer R, Pohl G, Adami S, Mautalen C, et al. Sustained vertebral fracture risk reduction after withdrawal of teriparatide in postmenopausal women with osteoporosis. *Arch Intern Med* 2004;164(18):2024-30.

[108] Davis S, Martyn-St James M, Sanderson J, Stevens J, Goka E, Rawdin A, et al. A systematic review and economic evaluation of bisphosphonates for the prevention of fragility fractures. *Health Technol Assess* 2016;20(78):1-406.

[109] Strom O, Landfeldt E, Garellick G. Residual effect after oral bisphosphonate treatment and healthy adherer effects--the Swedish Adherence Register Analysis (SARA). *Osteoporos Int* 2015;26(1):315-25.

[110] Tsourdi E, Langdahl B, Cohen-Solal M, Aubry-Rozier B, Eriksen EF, Guañabens N, et al. Discontinuation of denosumab therapy for osteoporosis: a systematic review and position statement by ECTS. *Bone* 2017;105:11-7.

[111] Miller PD, Bolognese MA, Lewiecki EM, McClung MR, Ding B, Austin M, et al. Effect of denosumab on bone density and turnover in postmenopausal women with low bone mass after long-term continued, discontinued, and restarting of therapy: a randomized blinded phase 2 clinical trial. *Bone* 2008;43(2):222-9.

[112] Davis S, Simpson E, Hamilton J, James MM, Rawdin A, Wong R, et al. Denosumab, raloxifene, romosozumab and teriparatide to prevent osteoporotic fragility fractures: a systematic review and economic evaluation. *Health Technology Assessment* 2020;24(29):1-314.

[113] Office for National Statistics (ONS). *National life tables, UK: 2012 to 2014 [Internet]*, 2020 [accessed 1.7.21] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables

[114] Strom O, Borgstrom F, Sen SS, Boonen S, Haentjens P, Johnell O, et al. Cost-effectiveness of alendronate in the treatment of postmenopausal women in 9 European countries--an economic evaluation based on the fracture intervention trial. *Osteoporos Int* 2007;18(8):1047-61.

[115] Office for National Statistics (ONS). *National life tables, UK: 2017 to 2019 [Internet]*, 2020 [accessed 28.9.21] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables

[116] Johnell O, Kanis JA, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Mortality after osteoporotic fractures. *Osteoporos Int* 2004;15(1):38-42.

[117] Barrett JA, Baron JA, Beach ML. Mortality and pulmonary embolism after fracture in the elderly. *Osteoporos Int* 2003;14(11):889-94.

[118] Oden A, Dawson A, Dere W, Johnell O, Jonsson B, Kanis JA. Lifetime risk of hip fractures is underestimated. *Osteoporos Int* 1998;8(6):599-603.

[119] Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B, Oglesby AK. The components of excess mortality after hip fracture. *Bone* 2003;32(5):468-473.

[120] Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone* 2004;35(2):375-82.

[121] National Institute for Health and Care Excellence. *Raloxifene and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE technology appraisal guidance 161 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta161

[122] National Institute for Health and Care Excellence. *Raloxifene for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE technology appraisal guidance 160 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta160

[123] National Institute for Health and Care Excellence. *Romosozumab for treating severe osteoporosis [ID3936]. Clarification letter. Single technology appraisal*. London: NICE, 2021

[124] Szende A, Janssen B, Cabases J. *Self-reported population health: an international perspective based on EQ-5D*. The Netherlands: Springer, 2014.

[125] Scottish Medicines Consortium. *Medicines advice romosozumab (Evenity)*, 2020

[126] Kanis JA, Johansson H, Oden A, Harvey NC, Gudnason V, Sanders KM, et al. Characteristics of recurrent fractures. *Osteoporos Int* 2018;29(8):1747-57.

[127] Adachi JD, Lyles KW, Colón-Emeric CS, Boonen S, Pieper CF, Mautalen C, et al. Zoledronic acid results in better health-related quality of life following hip fracture: the HORIZON-Recurrent Fracture Trial. *Osteoporos Int* 2011;22(9):2539-49.

[128] Blomfeldt R, Törnkvist H, Ponzer S, Söderqvist A, Tidermark J. Comparison of internal fixation with total hip replacement for displaced femoral neck fractures. Randomized, controlled trial performed at four years. *J Bone Joint Surg Am* 2005;87(8):1680-8.

[129] Ekström W, Németh G, Samnegård E, Dalen N, Tidermark J. Quality of life after a subtrochanteric fracture: a prospective cohort study on 87 elderly patients. *Injury* 2009;40(4):371-6.

[130] Lindgren P, Lindström J, Tuomilehto J, Uusitupa M, Peltonen M, Jönsson B, et al. Lifestyle intervention to prevent diabetes in men and women with impaired glucose tolerance is cost-effective. *Int J Technol Assess Health Care* 2007;23(2):177-83.

[131] Joint Formulary Committee. *British National Formulary [Internet]*. London: BMJ Group and Pharmaceutical Press, 2021 [accessed 1.7.21]. Available from: https://bnf.nice.org.uk/

[132] Eli Lilly and Company Limited. Forsteo 20 micrograms/80 microlitres solution for injection in pre-filled pen (teriparatide). Summary of product characteristics [Internet]. electronic Medicines Compendium (eMC), 2020 [accessed 1.7.21]. Available from: https://www.medicines.org.uk/emc/product/2215

[133] Curtis LA, Burns A. *Unit costs of health and social care 2020 [Internet]*. Canterbury: Personal Social Services Research Unit (PSSRU), University of Kent, 2020 [accessed 1.7.21] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

[134] NHS. *2020/21 National Tariff Payment System. The National Tariff Workbook 2020/21 [Internet]*: National Health Service, 2020 [accessed 1.7.21] Available from: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/

[135] Borgstrom F, Jonsson B, Strom O, Kanis JA. 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):1781-93.

[136] Gutierrez L, Roskell N, Castellsague J, Beard S, Rycroft C, Abeysinghe S, et al. Clinical burden and incremental cost of fractures in postmenopausal women in the United Kingdom. *Bone* 2012;51(3):324-31.

[137] Gutierrez L, Roskell N, Castellsague J, Beard S, Rycroft C, Abeysinghe S, et al. Study of the incremental cost and clinical burden of hip fractures in postmenopausal women in the United Kingdom. *J Med Econ* 2011;14(1):99-107.

[138] Office for National Statistics (ONS). Consumer price inflation tables [Internet]. ONS, 2021 [accessed 4.8.21]. Available from: https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation

[139] Ryder S, Fox K, Rane P, Armstrong N, Wei CY, Deshpande S, et al. A Systematic Review of Direct Cardiovascular Event Costs: An International Perspective. *Pharmacoeconomics* 2019;37(7):895-919.

[140] Danese MD, Gleeson M, Kutikova L, Griffiths RI, Azough A, Khunti K, et al. Estimating the economic burden of cardiovascular events in patients receiving lipid-modifying therapy in the UK. *BMJ OPEN* 2016;6(8):e011805.

[141] Chapman RH, Liu LZ, Girase PG, Straka RJ. Determining initial and follow-up costs of cardiovascular events in a US managed care population. *BMC Cardiovasc Disord* 2011;11:11.

[142] van Staa TP, Kanis JA, Geusens P, Boonen A, Leufkens HG, Cooper C. The cost-effectiveness of bisphosphonates in postmenopausal women based on individual long-term fracture risks. *Value Health* 2007;10(5):348-57.

[143] UCB Pharma Ltd. *UCB Data on File - National Institute of Health and Care Excellence scientific advice. Preliminary independent model advice + clarification questions. Cost-effectiveness model for romosozumab for the treatment of osteoporosis*, 2017

[144] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Romosozumab for treating severe osteoporosis [ID3936]

Produced by	Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Rob Riemsma, Reviews Manager, KSR Ltd, United Kingdom (UK) Pim Wetzelaer, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), EUR, the Netherlands Sean Harrison, Systematic Reviewer and Statistician, KSR Ltd, UK Susan O'Meara, Systematic Reviewer, KSR Ltd, UK Hannah Penton, Health Economics Researcher, ESHPM, EUR, the Netherlands Steve Ryder, Health Economist, KSR Ltd, UK Steven Duffy, Information Specialist, KSR Ltd, UK Nigel Armstrong, Health Economics Manager, KSR Ltd, UK Maiwenn Al, Health Economics Researcher, ESHPM, EUR, the Netherlands Jos Kleijnen, Director, KSR Ltd, UK
Correspondence to	Rob Riemsma, Kleijnen Systematic Reviews Ltd Unit 6, Escrick Business Park Riccall Road, Escrick York, YO19 6FD United Kingdom
Date completed	05/10/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number STA 13/51/35.

Declared competing interests of the authors

None.

Acknowledgements

None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout 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 Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Riemsma R, Wetzelaer P, Corro Ramos I, Harrison S, O'Meara S, Penton H, Ryder S, Duffy S, Armstrong N, Al M, Kleijnen J. Romosozumab for treating severe osteoporosis [ID3936]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors

Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Pim Wetzelaer acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Isaac Corro Ramos, Hannah Penton, Steve Ryder and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Sean Harrison acted as systematic reviewer and statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Susan O'Meara acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Maiwenn Al critiqued the company's economic evaluation, contributed to the writing of the report, and provided general health economic guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse event
AFF	Atypical femur fractures
AiC	Academic in confidence
ALN	Alendronate
ARR	Absolute risk reduction
ASBMR	American Society for Bone and Mineral Research
AUS	Australia
AWMSG	All Wales Medicines Strategy Group
β-CTX	Beta-C-Terminal Telopeptide of Type 1 Collagen
BC	Base-case
BMI	Body mass index
BMD	Bone mineral density
BNF	British National Formulary
BP	Bisphosphonates
BPI	Brief Pain Inventory
BTM	Bone turnover marker
CABG	Coronary artery bypass graft
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Central Register of Controlled Trials
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CG	Clinical guideline
CHD	Coronary heart disease
CHMP	Committee for Medicinal Products for Human Use
CiC	Commercial in confidence
CPRD	Clinical Practice Research data
CRD	Centre for Reviews and Dissemination
CPI	Consumer price index
Crl	Credible interval
CS	Company submission
CSR	Clinical study report
CV	Cardiovascular
DARE	Database of Abstracts of Reviews of Effects
DEN	Denosumab
DSA	Deterministic sensitivity analysis
DXA	Dual-energy x-ray absorptiometry
ECTS	European Calcified Tissue Society
EED	Economic Evaluation Database
EEPIA	European Federation of Pharmaceutical Industry Associations
EMA	European Medicines Agency
EQ-5D	EuroQoL-5 Dimensions
EQ-5D-3L	EuroQoL-5 Dimensions-3 Levels
EQ-5D-5L	EuroQoL-5 Dimensions-5 Levels
EQ-VAS	EuroQoL-Visual analogue scale
ERG	Evidence Review Group
ESCEO	European Society for Clinical and Economic Aspects of Osteoporosis
ESHPM	Erasmus School of Health Policy & Management
EU	European Union
EULAR	European League Against Rheumatism
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FFN	Fragility Fracture Network

FN	Femoral neck
FRAX	Fracture Risk Assessment tool
FSH	Follicle-stimulating hormone
GPRD	General Practice Research Database
GI	Gastrointestinal
GIAE	Gastrointestinal adverse event
GIN	Guidelines International Network
GP	General practitioner
HES	Hospital episodes statistics
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
ICUROS	International Costs and Utilities Related to Osteoporotic Fractures Study
iMTA	Institute for Medical Technology Assessment
INAHTA	International Network of Agencies for Health Technology Assessment
Inc.	Incremental
IOF	International Osteoporosis Foundation
IQR	Interquartile range
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITT	Intention-to-treat
IU	International unit
IV	Intravenous
IVRS	Interactive voice response system
KM	Kaplan Meier
KSR	Kleijnen Systematic Reviews
LAD	Limited activity days
LOCF	Last observation carried forward
LS	Lumbar spine
LTC	Long-term care
LYG	Life years gained
m	Months
MD	Mean difference
MeSH	Medical subject headings
MI	Myocardial infarction
MOF	Major osteoporotic fracture
MTA	Multiple technology appraisal
NA	Not applicable
NAm	North America
NCPE	National Centre for Pharmacoeconomics
NHNV	Non-hip, non-vertebral
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NLM	National Library of Medicine
NMA	Network meta-analysis
NOF	National Osteoporosis Foundation
NOGG	National Osteoporosis Guideline Group
NOS	National Osteoporosis Society
NR	Not reported
ONJ	Osteonecrosis of the jaw
ONS	Office for National Statistics
OPAQ-SV	Osteoporosis Assessment Questionnaire Short Version
OR	Odds ratio

OS	Overall survival
P1NP	Procollagen Type 1 N-Telopeptide
PAS	Patient access scheme
PASS	Post-Authorization Safety Studies
PICOS	Population, intervention, comparator, outcome, study design
PO	Oral administration
PRIMA	Preliminary independent model advice
PSA	Probability sensitivity analysis
PSP	Patient support programme
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTCA	Percutaneous transluminal coronary angioplasty
Q6M	Once every six months
QALY	Quality adjusted life year
QD	Once daily
QM	Once monthly
QoL	Quality of life
QW	Once weekly
r	Exposure-adjusted incidence rate per 100 subject-years
RAL	Raloxifene
RANK	Receptor activator of nuclear factor kappa-B
RCT	Randomised controlled trial
RIS	Risedronate
RoB	Risk of bias
ROM/ROMO	Romosozumab
RR	Risk ratio/relative risk
RRR	Relative risk reduction
SAE	Serious adverse event
SC	Subcutaneous
ScHARR	School of Health and Related Research
ScHARRHUD	School of Health and Related Research Health Utilities Database
SD	Standard deviation
SE	Standard error
SERM	Selective oestrogen receptor modulator
SI	Système international (d'unités); English: International System of Units
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SMR	Standardised mortality ratio
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TH	Total hip
TLV	Tandvårds- och läkemedelsförmånsverket; English: The Swedish Dental and Pharmaceutical Benefits Agency
TRP	Teriparatide
TTO	Time trade-off
Tx	Treatment
UK	United Kingdom
USA	United States of America
VBA	Visual Basic
Vert.	Vertebral
WCO-IOF-ESCEO	World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases
WHO	World Health Organization
WTP	Willingness-to-pay
ZOL	Zoledronate

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues, Section 1.2 presents the key model outcomes, Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness (CE). Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (CE) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1 provides a summary of the key issues identified by the ERG.

Table 1.1: Summary of key issues

ID3936	Summary of issue	Report Sections
1	There is a problem with the population in the CS, with comparator populations at different risks for fracture, which means none of the comparisons are reliable	2.1 and 3.4
2	It is possible that effects of romosozumab wane after 42 months	3.2.5 and 3.6
3	The network meta-analyses (NMAs) are unreliable	3.3 and 3.4
4	It is unclear whether the company's and ERG's base-case analyses are representative for UK clinical practice	4.2.4, 5.1 and 6.2
5	Assumptions regarding persistence with osteoporosis therapies are uncertain and have a major impact on the model results	4.2.6
6	Model usability could be improved by performing calculations in the model work sheets and by significantly reducing its running time	5.3

CS = company submission; ERG = Evidence Review Group; NMA = network meta-analysis; UK = United Kingdom

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are the following:

- Persistence rates for romosozumab and alendronate,
- Excess mortality associated to fractures (ERG assumed only for hip fractures and company also after vertebral and non-hip non-vertebral (NHNV) fractures),
- Incremental fracture and daily long-term care (LTC) costs,
- Inclusion of cardiovascular (CV) adverse events (AEs) and patient support programme (PSP) costs,
- Number of General Practitioner (GP) visits per year, and
- The source of United Kingdom (UK) general population mortality rates.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival; OS) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Reducing the incidence of fractures, and
- QALYs are reduced by cardiovascular (CV) adverse events (AEs).

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments, and
- Reducing costs associated to a decreased number of fractures.

The modelling assumptions that have the greatest effect on the ICER are:

- Treatment persistence
- Treatment effect of romosozumab followed by alendronate and alendronate alone
- Utility multipliers for hip, vertebral and NHNV fracture
- Comparator choice
- Inclusion of CV AEs
- Assumed excess mortality
- Start age of the population
- Model time horizon

1.3 The decision problem: summary of the ERG's key issues

The decision problem addressed in the company submission (CS) is broadly in line with the final scope issued by NICE. However, there is a problem with the population in the CS, which means none of the comparisons are reliable (Table 1.2).

Table 1.2: Key issue 1: There is a problem with the population in the CS, with comparator populations at different risks for fracture, which means none of the comparisons are reliable

Report Section	Sections 2.1 and 3.4
Description of issue and why the ERG has identified it as important	• The population in the CS (imminent risk of a fracture, i.e. having had a MOF within the last 2 years) is narrower than the scope, which does not define "high risk" or mention a time limit, and the ARCH ITT population where some patients without any time limit were included. In the NMAs the populations in the comparator studies are diverse, but mainly include women at high risk of a fracture as in the ARCH ITT population. • The ARCH trial includes a head-to-head comparison of romosozumab vs. alendronate. Both treatments are recommended for women at high risk of a fracture. However, oral bisphosphonates (such as alendronate) are recommended for the "high risk" group and anabolic agents (such as romosozumab) are recommended for the "very high risk" group (Kanis et al. 2020). Therefore, the comparison, romosozumab vs. alendronate may not be the appropriate comparison in the very high risk subgroup.

Report Section	Sections 2.1 and 3.4
What alternative approach has the ERG suggested?	The submission should only focus on the "imminent risk" population in the ARCH trial. This population is as specified in the CS and allows a head-to-head comparison with alendronate.
What is the expected effect on the cost effectiveness estimates?	The effectiveness results used in the model are based on the NMA for the ITT population. However, the overall model is based on a different population, the imminent risk population. It would be useful if the company could add a scenario where both effectiveness data and the whole model are based on the imminent risk population from the ARCH trial.
What additional evidence or analyses might help to resolve this key issue?	Clinical expert opinion as to whether imminent risk is consistent with only high as opposed to very high risk or whether it also includes very high risk. This would provide clarity as to whether alendronate is the most appropriate comparator.
CS = company submission; ERG = Evidence Review Group; ITT = intention-to-treat; MOF = major osteoporotic fracture; NMA = network meta-analysis	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The ERG identified two major concerns with the evidence presented on the clinical effectiveness, namely that it is possible that effects of romosozumab wane after 42 months (Table 1.3) and that the network meta-analyses (NMAs) are unreliable (Table 1.4).

Table 1.3: Key issue 2: It is possible that effects of romosozumab wane after 42 months

Report Section	Sections 3.2.5 and 3.6
Description of issue and why the ERG has identified it as important	The Kaplan-Meier curves for time to first clinical fracture and time to first non-vertebral fracture show that there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first fracture up to month 42. However, the curves seem to converge again by month 48. This means that it is possible that the effects of romosozumab wane over time. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Therefore, longer term follow-up is needed to see whether the effects are maintained over time.
What alternative approach has the ERG suggested?	The economic evaluation should include a scenario where treatment waning starts at 4 years followed by a dynamic offset (linear waning) of the treatment effect. The economic evaluation should also include a scenario where the dynamic offset of the treatment effect is shorter (e.g., three years).
What is the expected effect on the cost effectiveness estimates?	In the base-case analysis, treatment effect is maintained for 5 years (60 months). After that, a dynamic offset (linear waning) of the treatment effect is assumed for another 5 years. At year 11, there is no treatment effect. An early treatment effect waning can be modelled by using larger hazard ratios. This would increase the ICER.
What additional evidence or analyses might help to resolve this key issue?	For the first scenario no additional evidence is necessary. For the second scenario the company would need to adjust the model to allow selecting different durations of the dynamic offset of the treatment effect.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 1.4: Key issue 3: The network meta-analyses (NMAs) are unreliable

Report Section	Sections 3.3 and 3.4
Description of issue and why the ERG has identified it as important	The NMAs are unreliable for the following reasons: • There was little direct evidence for comparisons for romosozumab included in any of the NMAs. • Most studies had differences in mean age, ethnicity, or rate of prevalent vertebral fractures, indicating at least a moderate RoB from effect modification. • As almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone. • Individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). • There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the risk of bias. • As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low risk of bias; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high risk of bias.
What alternative approach has the ERG suggested?	There is no alternative approach with the data available in the CS, beyond interpreting the effect estimates with due caution from the high-RoB present in almost all comparisons, with the exceptions of alendronate and placebo (which had direct evidence). To reduce bias, either of the following is possible, though would require additional data: 1. Include direct evidence from more trials of romosozumab and comparator treatments, by conducting more trials; and 2. Request individual participant data from all trials included in the NMAs and adjust for known effect modifiers. This option only decreases bias, and large biases may remain due to an inability to adjust away all effects of differences in effect modifiers between trials.
What is the expected effect on the cost effectiveness estimates?	The expected effect on the CE estimates is uncertain.

Report Section	Sections 3.3 and 3.4
What additional evidence or analyses might help to resolve this key issue?	To reduce bias, either of the following is possible, though would require additional data: 1. Include direct evidence from more trials of romosozumab and comparator treatments, by conducting more trials; and 2. Request individual participant data from all trials included in the NMAs and adjust for known effect modifiers. This option only decreases bias, and large biases may remain due to an inability to adjust away all effects of differences in effect modifiers between trials.
CE = cost effectiveness; CS = company submission; ERG = Evidence Review Group; NMA = network meta-analysis; RoB = risk of bias	

1.5 The cost effectiveness evidence: summary of the ERG's key issues

A full summary of the CE evidence review conclusions can be found in Section 6.4 of this report. The company's CE results are presented in Section 5, the ERG's summary and detailed critique are in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues in the CE evidence are discussed in Tables 1.5 to 1.7.

Table 1.5: Key issue 4: It is unclear whether of the company's and ERG's base-case analyses are representative for UK clinical practice

Report Section	Sections 4.2.4, 5.1 and 6.2
Description of issue and why the ERG has identified it as important	There is uncertainty regarding the appropriateness and relevance of the comparators included in the analyses, and how these relate to the relevant population for this assessment as described in key issue 1. For example, Kanis et al. 2020 recommended that raloxifene is given to patients at low risk of fractures, oral bisphosphonates (such as alendronate and risedronate) are given to high risk patients, and anabolic agents (such as romosozumab and teriparatide) followed by an inhibitor of bone resorption (such as oral bisphosphonates) are provided to very high risk patients.
What alternative approach has the ERG suggested?	Identify what comparators are representative of UK clinical practice in the imminent risk population. After this is done, results can be selected for the right comparators only.
What is the expected effect on the cost effectiveness estimates?	As shown with the different scenario analyses, results are likely to vary depending on the comparators selected.
What additional evidence or analyses might help to resolve this key issue?	The Committee should clarify what comparators are representative of UK clinical practice in the imminent risk population.
ERG = Evidence Review Group; UK = United Kingdom	

Table 1.6: Key issue 5: Assumptions regarding persistence with osteoporosis therapies are uncertain and have a major impact on the model results

Report Section	Section 4.2.6
Description of issue and why the ERG has identified it as important	The company's approach to model persistence is inconsistent between intervention (persistence based on trial data) and comparators (persistence based on clinical practice) and is likely to be biased in favour of the intervention. Persistence assumptions

Report Section	Section 4.2.6
	were identified as one of the most important drivers of the CE results.
What alternative approach has the ERG suggested?	The ERG estimates for persistence are consistent between intervention and comparators. The ERG also identified a more recent study (Morley et al. 2020) to estimate persistence on the comparator treatments.
What is the expected effect on the cost effectiveness estimates?	When the ERG preferred base-case assumption for persistence with alendronate is applied (without the other ERG preferred changes) to the company base-case model, the ICER increased from £16,660 to £162,391 per QALY gained.
What additional evidence or analyses might help to resolve this key issue?	The uncertainty regarding persistence with osteoporosis therapies could be resolved by a study that uses data on present-day persistence in the UK, and by further investigating to what extent it is relevant to distinguish between naïve and non-naïve patients.
CE = cost effectiveness; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; UK = United Kingdom	

Table 1.7: Key issue 6: Model usability could be improved by performing calculations in the model work sheets and by significantly reducing its running time

Report Section	Section 5.3
Description of issue and why the ERG has identified it as important	Model review would be facilitated if calculations were performed in the model worksheets, instead of being hard coded in VBA. This code was initially password protected and therefore the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. After clarification, the company provided most of the VBA code which was reviewed by the ERG. No major issues were found but, nevertheless, the ERG was not allowed to make any changes to the VBA code in the model version used to run the scenarios because this model version still contains the code used for the Fracture Risk Assessment tool (FRAX), which is confidential. Additionally, the model seems to be extremely demanding regarding the computational power needed to run within a reasonable time. This makes the validation process extra difficult. The ERG did not succeed in running any probabilistic sensitivity analysis (PSA). Some counterintuitive results were observed when teriparatide was involved as a comparator treatment. The ERG was not able to find the source for these inconsistencies, which might need further confirmation from the company.
What alternative approach has the ERG suggested?	A full evaluation of the model and the assumptions included cannot be performed without access to the VBA code within the model. The ERG would like to suggest the company conduct an analysis to estimate the minimal PSA loop sizes that would provide reliable results in a minimum running time and to re-consider the programming of the model in order to make it computationally more efficient.
What is the expected effect on the cost effectiveness estimates?	It should not impact the model results but it would facilitate model validation and usability.

What additional evidence or analyses might help to resolve this key issue?	A new model version in which the ERG is allowed to make changes in the VBA code if deemed necessary. Also, a new model version with improved running time would enable the execution of a PSA.
ERG = Evidence Review Group, FRAX = Fracture Risk Assessment tool; PSA = probabilistic sensitivity analysis; VBA = Visual Basic	

1.6 Other key issues: summary of the ERG's view

No other key issues were identified by the ERG.

1.7 Summary of the ERG's view

Table 1.8 provides the incremental results of both the company's and ERG's preferred base-cases, as well as the impact of each ERG assumption change applied individually to the company base-case. As can be seen, the ERG base-case ICER is substantially larger than the company's. The change which had the largest impact by far on the results was the use of estimates for persistence on alendronate from Morley et al. 2020, which increased the ICER to £162,391. The next largest change in results was observed when assuming a daily cost of long-term care of £67 (i.e., instead of £112), which increased the ICER by nearly £6,000 per QALY gained. All other changes had an independent impact of less than £5,000 on the ICER.

The ERG was unable to run a probabilistic sensitivity analysis (PSA) for its preferred base-case analysis. However, given the deterministic ICER and assuming that the PSA ICER would be in line with this one, the probability that romosozumab is considered cost effective at a threshold of £30,000 compared to alendronate is likely to be [redacted]%. Scenario analyses run on the ERG preferred assumptions showed that model results were most sensitive to assumed rates of persistence; however, scenarios surrounding utility multipliers, treatment effect waning, excess mortality due to fractures and inclusion of CV AEs and PSP also had large impacts on the ICER, which was very sensitive to changes in the small incremental QALYs. When various alternative comparators were included in the analysis, romosozumab was dominated by zoledronate. In this situation, the only relevant comparison was zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated.

Table 1.8: Summary of ERG's preferred assumptions and ICER

Scenario	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
Company's base-case	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	162,391
+ Excess mortality only for hip fractures	[redacted]	[redacted]	17,185
+ Daily LTC costs £67	[redacted]	[redacted]	22,476
+ Incremental fracture costs	[redacted]	[redacted]	20,398
+ CV adverse events included	[redacted]	[redacted]	19,500
+ No PSP	[redacted]	[redacted]	17,680
+ 2 GP visits per year	[redacted]	[redacted]	17,117
+ UK general population mortality 2017-2019	[redacted]	[redacted]	16,903
ERG's preferred base-case	[redacted]	[redacted]	483,750

Scenario	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
CV = cardiovascular; ERG = Evidence Review Group; GP = General Practitioner; ICER = incremental cost effectiveness ratio; LTC = long-term care; PSP = probabilistic sensitivity analysis; QALY = quality adjusted life year; UK = United Kingdom			

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Postmenopausal women with severe osteoporosis at high risk of fracture	Postmenopausal women with severe osteoporosis who are at high risk of fracture and who have: • Experienced a recent MOF within 24 months; and • Thus, are at imminent risk of another fragility fracture	• Romosozumab is not licensed for use in men, in premenopausal women or in patients without severe osteoporosis • The submission positions romosozumab for use in a population that is part of the licenced population, including women with the greatest unmet need, and for whom romosozumab is expected to provide substantial clinical benefit	The population is not in line with the NICE scope. The population described in the NICE scope is the same as the licensed population for romosozumab. However, the population in the ARCH trial is narrower in that patients should have had a previous MOF. The population in the CS is narrower again in that a patient should have had a recent (within 24 months) MOF.
Intervention	Romosozumab	Romosozumab for 12 months, followed by sequential alendronate.	Romosozumab is licensed as a 12-month course of treatment. The SmPC for romosozumab states that *"following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months"*	The intervention in the CS is romosozumab for 12 months, followed by sequential alendronate.
Comparator(s)	• Bisphosphonates (alendronic acid, risedronate sodium, ibandronic acid and zoledronic acid)	The base-case comparisons are vs. alendronate, using the head-to-head ARCH study, and vs. no active treatment. Scenario analyses are provided against all other comparators	No trials of the licensed dose of ibandronate were found to be included in the NMA for fracture outcomes, therefore comparisons could not be conducted.	The comparators are in line with the NICE scope, except for the exclusion of ibandronate.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• Non-bisphosphonates (denosumab, raloxifene and teriparatide) • No active treatment	listed in the scope, using the NMA, except ibandronic acid.		
Outcomes	• Osteoporotic fragility fracture • Bone mineral density • Mortality • Adverse effects of treatment • Health-related quality of life	In line with the final NICE scope.	In line with the final NICE scope.	The outcomes reported are in line with the NICE scope.
Economic analysis	• The reference case stipulates that the CE of treatments should be expressed in terms of incremental cost per QALY • The reference case stipulates that the time horizon for estimating clinical and CE should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared • Costs will be considered from an NHS and PSS perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent	Not reported.	Not reported.	The CE analyses were conducted according to the NICE reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	treatment technologies will be taken into account			
Subgroups to be considered	If evidence allows, subgroups based on patient characteristics that increase the risk of fracture (that is, those specified in NICE clinical guideline 146) or that affect the impact of fracture on lifetime costs and outcomes should be considered.	Not reported.	Not reported.	No subgroup analyses were performed by the company.

Based on Table 1 and pages 11 to 12 of the CS[1]

CE = cost effectiveness; CS = company submission; ERG = Evidence Review Group; MOF = major osteoporotic fracture; NICE = National Institute for Health and Care Excellence; NMA = network meta-analysis; NHS = National Health Service; PSS = Personal Social Services; QALY = quality adjusted life year; SmPC = Summary of Product Characteristics

2.1 Population

The population defined in the scope is: "*Postmenopausal women with severe osteoporosis at high risk of fracture"*.[2]

The population in the company submission (CS) is limited to *"Postmenopausal women with severe osteoporosis who are at high risk of fracture and who have: Experienced a recent major osteoporotic fracture (MOF) within 24 months; and thus, are at imminent risk of another fragility fracture"*.[1]

According to the company, the decision problem addressed in the CS is narrower than that specified in the final scope and narrower than the marketing authorisation for romosozumab (CS, Section B.1.1, page 10).[1] According to the company, the patient population in the CS *"focusses on women with the greatest unmet need, and for whom romosozumab is expected to provide substantial clinical benefit"* (CS, Section B.1.1, page 10).[1]

The population included in the ARCH trial was ambulatory postmenopausal women aged 55 to 90 years if they had at least one of the following bone mineral density (BMD) and fracture criteria:

- BMD T-score at the total hip or femoral neck of ≤ -2.50 and EITHER:
    - at least one moderate (SQ2) or severe (SQ3) vertebral fracture OR
    - at least two mild (SQ1) vertebral fractures

OR

- BMD T-score at the total hip or femoral neck of ≤ -2.00 and EITHER:
    - at least two moderate (SQ2) or severe (SQ3) vertebral fractures OR
    - a fracture of the proximal femur that occurred within three to 24 months prior to randomisation

In addition, at least one hip must have been evaluable by dual-energy x-ray absorptiometry (DXA).

Assuming that all vertebral fractures are considered major osteoporotic fractures (MOFs), the population in the CS is largely in line with the population in the main trial, the ARCH trial, in which postmenopausal women who have previously suffered a MOF have been included.[3] However, the company does explain that the ARCH population is not completely in line with the population in the CS, with the key difference being that the ARCH trial did not mandate the prior fracture to be recent, whereas the romosozumab target population (i.e. the population in the CS) defines recency of fracture as a criterion (CS, page 43).[1]

In the ARCH trial, a total of [redacted] patients had suffered a fracture within zero to 24 months before randomisation ([redacted] in the romosozumab/alendronate group; [redacted] in the alendronate alone group). Of these, [redacted] patients in the romosozumab/alendronate group and [redacted] patients in the alendronate alone group suffered a recent MOF and would be eligible for treatment with romosozumab according to the target patient population considered in the CS.

In 2019, a European marketing authorisation was granted for romosozumab. Romosozumab is indicated for the treatment of severe osteoporosis in postmenopausal women at high risk of fracture.[4] Romosozumab is contraindicated for patients with: hypersensitivity to the active substance(s) or to any of the excipients, hypocalcaemia, or a history of MI or stroke.[4]

In summary, there seem to be three relevant populations:

1. The population as described in the NICE final scope,[2] which is the same as the European marketing authorisation for romosozumab: Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is not defined;
2. The population in the ARCH trial (intention-to-treat (ITT) population):[3] Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is defined as having previously suffered a MOF; and
3. The population in the CS:[1] Postmenopausal women with severe osteoporosis at high risk of fracture; where 'high risk of fracture' is defined as having suffered a fracture within the last two years (also referred to as 'imminent risk of fracture').

There is also a lack of clarity as to the difference between "high risk" and "very high risk". For example, Kanis et al. 2020 recommended that raloxifene is given to patients at low risk of fractures, oral bisphosphonates (such as alendronate and risedronate), are given to high risk patients, and anabolic agents (such as romosozumab and teriparatide) followed by an inhibitor of bone resorption (such as oral bisphosphonates) are provided to very high risk patients. However, it is not clear whether current clinical practice in the UK is based on these or similar recommendations. Multiple treatment guidelines are available that differ in their (wording of) recommendations and it is not clear which treatment guideline is both up-to-date and relevant for the NHS. This therefore raises the question as to whether "high" and "very high" are mutually exclusive or whether "high" includes "very high": if the former, then comparators other than alendronate might not be appropriate comparators, but if the latter then they might be.

2.2 Intervention

The intervention (romosozumab) is in line with the scope. However, romosozumab is licensed as a 12-month course of treatment. The Summary of Product Characteristics (SmPC) for romosozumab states that *"following completion of romosozumab therapy, transition to antiresorptive therapy is recommended in order to extend the benefit achieved with romosozumab beyond 12 months"*.[4] Therefore, the intervention in the CS is *"romosozumab for 12 months, followed by sequential alendronate"* (CS, Table 1, page 11).[1]

The recommended dose of romosozumab is 210 mg, which is administered as two subcutaneous (SC) injections of 105 mg each into the abdomen, thigh or upper arm.[4] The use of romosozumab is limited to once during a lifetime (CS, page 22).[1]

According to the company, no additional tests or investigations are required prior to the administration of romosozumab (CS, page 13).[1]

2.3 Comparators

The description of the comparators in the NICE scope is as follows: *"Bisphosphonates (alendronic acid, risedronate sodium, ibandronic acid and zoledronic acid), Non-bisphosphonates (including antiresorptive agents (denosumab, raloxifene) and anabolic agents (teriparatide)), and No active treatment"*.[2]

In the CS, the base-case comparisons are vs. alendronate, using the head-to-head ARCH study, and vs. no active treatment. Scenario analyses are provided against all other comparators listed in the scope, using the network meta-analysis (NMA),, except ibandronic acid. According to the company, *"no trials*

of the licensed dose of ibandronate were found to be included in the NMA for fracture outcomes, therefore comparisons could not be conducted" (CS, Table 1, page 11).[1]

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- Osteoporotic fragility fracture
- Bone mineral density
- Mortality
- Adverse effects of treatment
- Health-related quality of life

These were all assessed in the ARCH trial. However, the ARCH trial had a median follow-up duration of 33 months, at which time 90 participants in each group had died.[3] Therefore, if romosozumab is expected to improve survival, the follow-up is insufficient to show any differences.

Regarding health-related quality of life (HRQoL), the company states that the trial data do not provide HRQoL values sensitive to decreases in HRQoL after a fracture. In addition, the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited, according to the company.[1]

2.5 Other relevant factors

According to the company, romosozumab is innovative because it *"is the only dual-acting osteoanabolic biologic, with all other treatments being antiresorptives or a single-action anabolic. Antiresorptive therapies do not directly stimulate bone formation and therefore, romosozumab provides a clear advantage over bisphosphonates by rapidly increasing bone formation on naïve bone surface resulting in rapid improvements in bone density, mass, microstructure and strength leading to superior fracture risk reductions"* [5, 6] (CS, Section B.2.11).[1]

A patient access scheme (PAS) has been proposed for romosozumab. The proposed romosozumab with PAS price is £[redacted] per monthly dose, equivalent to a percentage discount of [redacted]%. This equates to an annual cost of £[redacted] (with PAS; CS, Section B.1.2, page 13).[1]

This appraisal does not fulfil the end-of-life criteria as specified by NICE because the life expectancy of patients eligible for romosozumab is well beyond 24 months. Therefore, treatment is not indicated for patients with a short life expectancy (normally less than 24 months).

According to the company, romosozumab is only licensed for use in postmenopausal women, not men. However, *"osteoporosis is four times more likely to occur in women than men, and is prevalent in 21.8% of women (versus 6.8% of men) over the age of 50 in the UK"*[7] (CS, Section B.1.4).[1]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company conducted a systematic review (an original review and two updates) to evaluate the evidence on clinical effectiveness (efficacy and safety) of romosozumab for the treatment of osteoporosis in postmenopausal women at increased risk of fracture.[8] Section 3.1 critiques the methods of the review including: the search strategy; study inclusion criteria; data extraction; assessment of risk of bias; and data synthesis.

3.1.1 Searches

Appendix D of the CS provided details of the systematic literature searches used to identify clinical efficacy and safety evidence.[8] Database searches were conducted in August 2016, updated in March 2018, and updated again in September 2020. Summaries of the resources searched for each set of searches are provided in Tables 3.1 to 3.3.

Table 3.1: Resources searched for clinical efficacy and safety, August 2016

Resource	Host/source	Date range	Date searched
Databases			
Embase	OvidSP	1974 to 17 August 2016	18 August 2016
MEDLINE	OvidSP	1946 to August Week 2 2016	24 August 2016
MEDLINE In-Process Citations, Epub Ahead of Print & Daily Update	OvidSP	up to 24 August 2016	24 August 2016
PubMed	NLM	up to 25 August 2016	25 August 2016
CDSR	Wiley Online Library	Issue 8/August 2016	16 August 2016
CENTRAL	Wiley Online Library	Issue 7/July 2016	16 August 2016
DARE	Wiley Online Library	Issue 2/April 2015	16 August 2016
HTA Database	Wiley Online Library	Issue 3/July 2016	16 August 2016
PROSPERO	http://www.crd.york.ac.uk/PROSPERO/	Not reported	Not reported

Resource	Host/source	Date range	Date searched
GIN Library	http://www.g-i-n.net	Not reported	Not reported
Clinical Trial Registries			
ClinicalTrials.gov	https://clinicaltrials.gov	Not reported	Not reported
WHO ICTRP	http://www.who.int/ictrp/en	Not reported	Not reported
Conference proceedings			
NOF	https://www.nof.org/	2013 and 2014	26 August 2016
NOS	https://nos.org.uk/	2014	6 October 2016
WCO-IOF-ESCEO	http://www.wco-iof-esceo.org/	2013, 2014, 2015 and 2016	25 August 2016
HTA websites			
CADTH	https://www.cadth.ca/	Not reported	Not reported
EMA / CHMP	http://www.ema.europa.eu	Not reported	Not reported
NICE	http://www.nice.org.uk	Not reported	Not reported
NIHR	http://www.nets.nihr.ac.uk/	Not reported	Not reported
US Drugs @ FDA	https://www.accessdata.fda.gov/scripts/cder/daf/	Not reported	Not reported

CADTH = Canadian Agency for Drugs and Technologies in Health; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CHMP = Committee for Medicinal Products for Human Use; DARE = Database of Abstracts of Reviews of Effects; EMA = European Medicines Agency; FDA = Food & Drug Administration; GIN = Guidelines International Network; HTA = health technology assessment; ICTRP = International Clinical Trials Registry Platform; NICE = National Institute for Health and Care Excellence; NIHR = National Institute for Health Research; NLM = National Library of Medicine; NOF = National Osteoporosis Foundation; NOS = National Osteoporosis Society; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases; WHO = World Health Organization

Table 3.2: Resources searched for clinical efficacy and safety, March 2018

Resource	Host/source	Date range	Date searched
Databases			
Embase	OvidSP	1974 to 27 March 2018	28 March 2018

Resource	Host/source	Date range	Date searched
MEDLINE	OvidSP	1946 to March Week 3 2018	28 March 2018
MEDLINE In-Process Citations, Epub Ahead of Print & Daily Update	OvidSP	up to 27 March 2018	27 March 2018
PubMed	NLM	up to 28 March 2018	28 March 2018
CENTRAL	Wiley Online Library	Issue 12/ February 2018	28 March 2018
Northern Light Life Sciences Conference Abstracts	Ovid	2010 to Week 11 2018	Not reported
Clinical Trial Registries			
ClinicalTrials.gov	https://clinicaltrials.gov	Not reported	Not reported
WHO ICTRP	http://www.who.int/ictrp/en	Not reported	Not reported
Conference proceedings			
NOF	https://www.nof.org/	2013 to 2016	Not reported
NOS	https://nos.org.uk/	2014 and 2016	Not reported
WCO-IOF-ESCEO	http://www.wco-iof-esceo.org/	2013 to 2017	Not reported

CENTRAL = Cochrane Central Register of Controlled Trials; ICTRP = International Clinical Trials Registry Platform; NOF = National Osteoporosis Foundation; NOS = National Osteoporosis Society; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases; WHO = World Health Organization

Table 3.3: Resources searched for clinical efficacy and safety, September 2020

Resource	Host/source	Date range	Date searched
Databases			
Embase	Not reported	Not reported	Not reported
PubMed	Not reported	Not reported	Not reported
Cochrane	Not reported	Not reported	Not reported

ERG comment:

- The selection of databases searched was very comprehensive. Full details of the database searches including the database name, host platform and date searched, were provided.
- Conference proceedings were searched. Details of the conferences searched, URLs, and the date of the searches were provided. The search strategies or search terms used, and results were not reported in the CS.[1] In response to the request for clarification, the company explained that relevant

conference publications were identified from the Embase search and that an additional search for conference publications was conducted in Northern Light Life Sciences Conference Abstracts.[9] The search strategy used to search Northern Light Life Sciences Conference Abstracts was provided in response to the request for clarification.

- Trials registers were searched, but details of the search strategies or search terms used, dates of searches, and results were not reported in the CS. Details of the trials registers searched and the search strategies used were provided in response to the request for clarification.[9]
- Health technology assessment (HTA) organisation websites were searched, but details of the search terms used, dates of searches, and results were not reported in the CS.[1] Details of the search terms used were provided in response to the request for clarification.[9]
- Extensive use of truncation, proximity operators, synonyms, and subject headings (MeSH and EMTREE) were included in the search strategies. Cited study design search filters for randomised controlled trials (RCTs) were included. There were no language or date limits.
- Separate searches for safety data were not conducted. Ideally, a search for AEs should be carried out alongside the search for effectiveness.[10]
- Update searches were conducted in March 2018 and September 2020. Full details of the March 2018 searches were provided, but only the databases searched were provided for the September 2020 update. Details of the search strategies and results for the September 2020 update were provided in response to the request for clarification.[9] The September 2020 searches did not directly replicate the original 2016 and March 2018 searches.

3.1.2 Inclusion criteria

As stated above, the company performed a systematic review to evaluate the evidence on clinical effectiveness (efficacy and safety) of romosozumab for the treatment of osteoporosis in postmenopausal women at increased risk of fracture.[1] The original systematic review was conducted in 2016 and the two subsequent updates in 2018 and 2020.[8] The study eligibility criteria for the original and updated systematic reviews are summarised in Table 3.4 below.

Table 3.4: Eligibility criteria used in the original and updated systematic reviews of clinical effectiveness evidence

	Inclusion criteria	Exclusion criteria
Population	Studies had to include: • Postmenopausal women with osteoporosis at increased risk of fracture Where trials included a mixed population of participants where not all these inclusion criteria were fulfilled, the study was excluded unless separate data were reported for the population of interest.	Studies recruiting the following were excluded: • Women being studied for the prevention or treatment of glucocorticoid induced osteoporosis • Women with normal or unspecified BMD who have not been selected based on the presence of risk factors • Women with other indications for osteoporosis treatment e.g., Paget's disease, hypercalcaemia of malignancy, metastatic breast cancer
Interventions	The intervention of interest was romosozumab (CDP7851/AMG 785; Amgen Inc. and UCB Inc.), a monoclonal antibody that binds and	Not applicable.

	Inclusion criteria	Exclusion criteria
	inhibits sclerostin, a negative regulator of bone formation, dosed at 210 mg SC QM for 12 months for the treatment of osteoporosis.	

Comparators	Eligible comparator therapies were pharmacological therapies and those in development (in accordance with the UK, European, and US licensed indications):
	• Placebo (in accordance with NICE TAG4627)
	• Usual care e.g., vitamin D and calcium supplementation (in accordance with NICE TAG4627)
	• Antibody-based RANK ligand therapy:
	  • Denosumab (Prolia, AMG 162; Amgen Inc.)
	• Parathyroid hormone-based therapy:
	  • Teriparatide (Forteo/Forsteo; Eli Lily)
	  • Abaloparatide (BA058; Radius Health)
	• Bisphosphonates (in accordance with NICE TAG4627):
	  • Alendronate (Fosamax; Merck Sharp & Dohme; also available non-proprietary)
	  • Risedronate (Actonel; Procter & Gamble UK)
	  • Ibandronate (Boniva; Hoffman La Roche)
	  • Zoledronic acid/zoledronate (Aclasta/Reclast; Novartis)
	• Selective oestrogen receptor modulators (SERMs):
	  • Raloxifene (Evista, LY139481; Eli Lilly)
	  • Strontium ranelate (Protelos; Servier Laboratories) (*subsequently excluded)*
		The following interventions were excluded:
		• Odanacatib (Merck) – following September 2016 protocol amendment to inclusion criteria
		• Strontium ranelate (Protelos; Servier Laboratories) – following March 2018 protocol amendment to the inclusion criteria[a]
		• Combination therapies (with the exception of usual care as described above)
		• Interventions which were not administered in accordance with licensed indications
		• Interventions which were co-administered with any other therapy with the potential to augment bone unless concomitant treatments were specified in the SmPC and applied equivalently in all study arms.

Outcomes	Studies had to report the occurrence of at least one of the following fracture outcomes: • New vertebral fracture	Studies were excluded from the review if they: • Did not report at least one prespecified fracture outcome

	Inclusion criteria	Exclusion criteria
	• Clinical vertebral fracture • Non-vertebral fracture • Clinical fracture • Hip fracture Fracture outcomes were classified using the definitions provided in each specific study.	• Only reported fractures as part of the adverse event monitoring process (e.g., a BMD outcome study reporting fractures outcomes as adverse events was excluded) • Reported outcomes relating to fractures associated with major trauma (e.g., road traffic accidents). Studies that reported mixed trauma and/or non-trauma fracture, were only included if they reported separate data for relevant non-trauma fractures
Study design	To be included in the review, trials had to fulfil the following criteria:[b] • Use a parallel RCT design. This included randomised dose finding and formulation trials with either a placebo or active control arm and was not limited by study phase • Followed-up patients for at least 12 months	The following were excluded: • Systematic reviews and pooled analyses (used for reference checking purposes only and not included in the review, unless the data are not available from publications of the individual trials) • Studies based on animal models • Pre-clinical and biological studies • Narrative reviews, letters, editorials, and opinions
Language restrictions	• No restrictions for clinical effectiveness review. • English language only for review of economic evaluations, cost and resource use studies.	

Based on Table 13 of Appendix D of the CS.[8]
a Only relevant to the review update. b This was in accordance with relevant criteria from the recent HTA undertaken by NICE (ScHARR, The University of Sheffield) in March 2015 to assess TA464 - Bisphosphonates for prevention osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161).[11]
BMD = bone mineral density; CS = company submission; HTA = Health Technology Assessment; NICE = National Institute of Health and Care Excellence; QM = once monthly; RANK = receptor activator of nuclear factor kappa-B; RCT = randomised controlled trial; SC = subcutaneous; ScHARR = School of Health and Related Research; SERM = selective oestrogen receptor modulator; SmPC = Summary of Product Characteristics; TA = technology appraisal; UK = United Kingdom; US = United States (of America)

ERG comment:

Population

As outlined in Section 2.1, three relevant populations have been described. One of these is the ITT population in the ARCH trial (postmenopausal women with severe osteoporosis at high risk of fracture, the latter being defined as a previous MOF) which is used as the basis for a series of NMAs and economic modelling in the CS.[3]

We note that some placebo-controlled RCTs providing comparator arms for the NMAs recruit populations with different characteristics to those described in the ARCH trial[3] i.e., they recruit a proportion of participants without evidence of prevalent vertebral fracture at baseline. Examples include (with percentages indicating the proportion of women without prevalent vertebral fracture at baseline): two RCTs evaluating zoledronic acid (36% to 40%);[12, 13] one RCT evaluating raloxifene (75%);[14] and one RCT assessing denosumab (73%)[15]. These RCTs did not provide outcome data on subgroups defined according to presence/absence of prevalent vertebral fracture at baseline. Whilst the proportions with and without fracture at baseline were balanced across treatment groups within the individual RCTs, the populations were unlikely to be comparable to that of the ARCH trial in the context of NMA.[3]

Language restrictions

There were no language restrictions for the clinical effectiveness review and this is line with recommended good practice in SLRs.[16]

3.1.3 Critique of data extraction

In section D.2 of Appendix D of the CS, it is stated that data from each included trial were extracted into a Microsoft Excel template by a reviewer who was familiar with the subject area and validated by a second, independent reviewer.[8] The response to the clarification questions confirmed that disagreements were resolved through discussion and if necessary, by consulting a third reviewer.[9] Recommended good practice is dual, independent data extraction, particularly for outcome data.[16] In light of this, the possibility of errors within the data extraction cannot be discounted.

3.1.4 Quality assessment

Section D.2 of Appendix D explains that the risk of bias (RoB) within each included study was assessed using the Cochrane RoB tool for RCTs[8] and the company's response to the clarification questions confirmed that the original version of the tool was used.[9] Although this tool is appropriate for assessing the quality of RCTs, it is not clear why the most recent version was not used (Cochrane RoB 2).[16] One reviewer assessed the RoB and a second reviewer independently checked the assessment. Any discrepancies were resolved through consensus.[8]

3.1.5 Evidence synthesis

It was not feasible to pool the identified, eligible RCTs using direct data, pairwise meta-analysis because of differences in populations and treatment comparisons. An indirect treatment comparison was performed and this is discussed in Sections 3.3 and 3.4.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The clinical effectiveness evidence for romosozumab in severe osteoporosis in the CS is mainly based on the ARCH trial. Two other phase III clinical trials, the FRAME and STRUCTURE trials are mentioned in the CS as well. However, neither the FRAME nor STRUCTURE trial studied a patient population aligned to where the company expects romosozumab to be used in NHS clinical practice; therefore these two trials will be briefly discussed in Section 3.2.7 of this report.[17, 18] A fourth study, the BRIDGE study,[19] considered use in men, which is not part of the marketing authorisation for romosozumab; as such, no clinical effectiveness results are presented from BRIDGE in the CS. However, some data from BRIDGE are introduced in the safety section of the CS and will be discussed in Section 3.2.6 of this report.

3.2.1 Details of the included trial: the ARCH trial

The ARCH trial is a phase III, multicentre, randomised, double-blind trial, comparing romosozumab followed by alendronate vs. alendronate alone in postmenopausal women with severe osteoporosis and a fragility fracture (see Table 3.5).[3] This trial provides evidence for romosozumab in its expected position in the clinical pathway: a first-line therapy in patients who have previously suffered a MOF. Efficacy outcomes reported in the ARCH trial include incidence of clinical, vertebral, non-vertebral and hip fracture and percentage change from baseline in BMD. Data from the ARCH trial were used as the main data for the economic modelling in this submission.

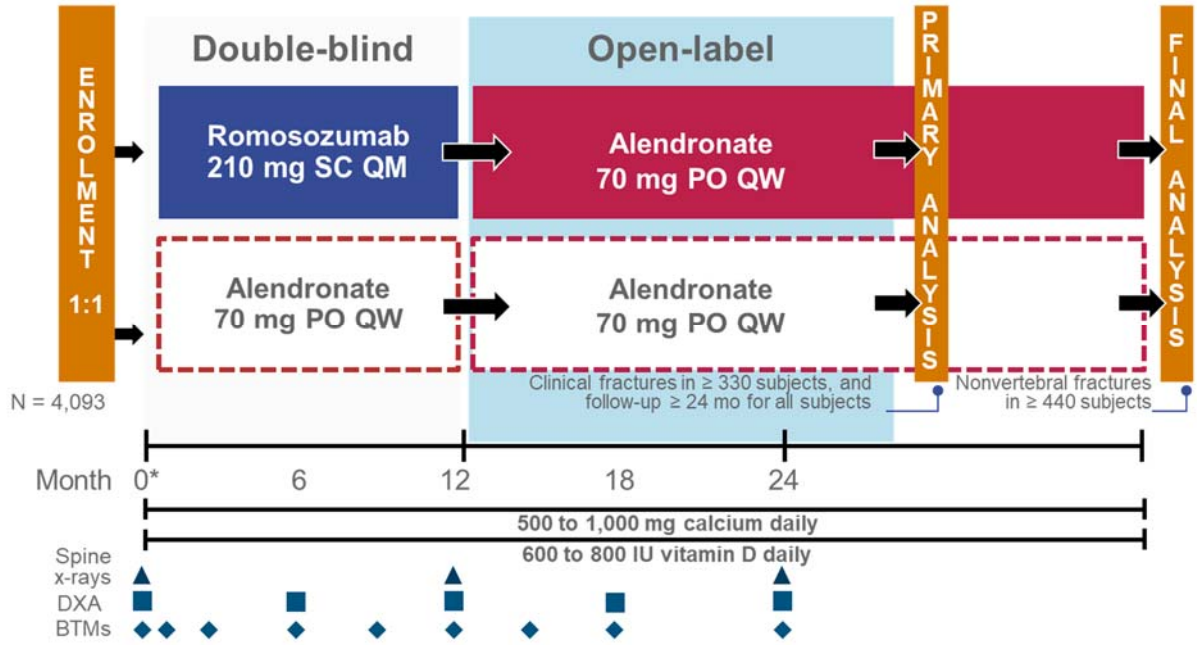
Table 3.5: Summary of methodologies for the ARCH trial

Trial number (acronym)	**NCT01631214 (ARCH)**
Study design	International, multicentre, randomised, double-blind, active-controlled, parallel-group, phase III.
Location	This study was conducted at [redacted] centres across Europe, North America, Central and South America, and Asia/Pacific, including [redacted] sites in the UK (76 UK patients out of 4,093).
Population	Ambulatory postmenopausal women aged ≥55 to ≤90 years of age at randomisation who met at least one of the following criteria: • BMD T-score of ≤–2.5 at TH or FN and either ≥1 moderate or severe vertebral fractures or ≥2 mild vertebral fractures • BMD T-score of ≤–2.0 at TH or FN and either ≥2 moderate or severe vertebral fractures, or a fracture of the proximal femur sustained three to 24 months prior to randomisation • At least one hip that could be evaluated by DXA
Duration of study	Double-blind treatment period: 12 months. Open-label period: minimum 12 months (until end of study).
Method of randomisation	Patients were randomly assigned to receive romosozumab or alendronate using IVRS. Randomisation was stratified by age (<75 years vs. ≥75 years).
Method of blinding	Double blind: patients and site staff remained blinded to the patient's original treatment assignment. Treatment assignment was only unblinded if the knowledge of the treatment was essential for the patient's further management.
Intervention(s)	Romosozumab (210 mg) QM SC for 12 months followed by open-label oral alendronate (70 mg) QW for at least 12 months (until study end).
Comparator(s)	Oral alendronate (70 mg) QW for 12 months followed by open-label alendronate (70 mg) for at least 12 months (until study end).
Permitted and disallowed concomitant medication	With the exception of the medications listed in the protocol, investigators may have prescribed any concomitant medications or treatments necessary to provide adequate supportive care.
Reported outcomes relevant to the decision problem	• Cumulative incidence of new vertebral fracture • Cumulative incidence of clinical fracture • Incidence of fractures (non-vertebral, all fractures, new or worsening vertebral, major non-vertebral, hip, MOF) • Percent change in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs

Trial number (acronym)	NCT01631214 (ARCH)
Based on CS, Tables 4 to 6, pages 29-33.[1] AE = adverse event; BMD = bone mineral density; BPI = Brief Pain Inventory; CS = company submission; DXA = dual-energy X-ray absorptiometry; EQ-5D = EuroQoL-5 Dimensions; FN = femoral neck; IVRS = interactive voice response system; LAD = limited activity days; LS = lumbar spine; MOF = major osteoporotic fracture; OPAQ-SV = Osteoporosis Assessment Questionnaire Short Version; QM = once monthly; QW = once weekly; SC = subcutaneous; TH = total hip; UK = United Kingdom	

The ARCH trial comprised the following study periods: initial screening and enrolment, double-blind treatment period, and open-label treatment period (Figure 3.1). Eligible patients were randomly assigned 1:1 to receive SC romosozumab 210 mg QM or oral alendronate 70 mg QW for the first 12 months (the double-blind period). Following this, patients received open-label oral alendronate 70 mg QW for the remainder of the study (the open-label period). Initial study drug given remained blinded until completion of the open-label period.

Figure 3.1: ARCH trial design



Based on CS, Figure 3, page 31.[1]

Footnotes: All patients received daily calcium (500 mg to 1,000 mg) and vitamin D (600 IU to 800 IU). *Patients with serum 25 (OH) vitamin D levels of ≥20 mg/mL and ≤40 ng/mL at screening received an initial loading dose of 50,000 to 60,000 IU of vitamin D. The final analysis (end-of-study) occurred when non-vertebral fracture events were confirmed for at least 440 subjects, or earlier if the primary analysis demonstrated superiority of romosozumab treatment for non-vertebral fracture risk reduction.

BTM = bone turnover markers; CS = company submission; DXA = dual-energy X-ray absorptiometry; IU = international unit; PO = oral administration; QM = once monthly; QW = once weekly; SC = subcutaneous

The ARCH trial was designed as an event-driven trial. The primary analysis for ARCH was performed after all patients had completed their month 24 visit and at least 330 patients had confirmed events of clinical fracture (composite of non-vertebral fracture and clinical vertebral fracture (a suspected vertebral fracture that is brought to medical attention and confirmed)). The median follow-up time at primary analysis was 2.7 years (33 months; interquartile range (IQR), 2.2 to 3.3). For all patients, BMD

was assessed at baseline and every 12 months at the lumbar spine, total hip and femoral neck by dual-energy X-ray absorptiometry (DXA).

The primary endpoints in the ARCH trial were the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis. Key secondary endpoints included incidence of non-vertebral fracture at primary analysis and percent change in BMD compared to baseline at months 12 and 24, at the lumbar spine, total hip, and femoral neck. Additional secondary endpoints included other fractures including hip fracture.

3.2.2 Statistical analyses of the ARCH trial

In the ARCH trial, a total of 4,093 patients were randomised to the initial treatment period, with 3,654 (89.3%) patients that completed the trial up to month 12 and 3,150 (77.0%) completed the primary analysis period. The trial population used for the analysis of outcomes in ARCH are detailed in Table 3.6.

Table 3.6: Trial populations for the ARCH trial

Analysis	NCT01631214 (ARCH)
Per protocol analysis set	Included patients in the full analysis set (for clinical and non-vertebral fracture) and the primary efficacy analysis set for vertebral fractures (for new vertebral fractures) who received active investigational products and met all of the patient eligibility criteria. Used to analyse clinical fracture, new vertebral fracture, and non-vertebral fracture through month 24, clinical and non-vertebral fracture at time of primary analysis, and non-vertebral fracture at final analysis as a sensitivity analysis.
Full analysis set	Included all randomised patients in the trial. They were analysed according to their randomised treatment assignments. This was the primary analysis set used for non-vertebral fracture, clinical fracture, clinical vertebral fracture, all fracture, major non-vertebral fracture, MOF, and hip fracture endpoints.
Primary efficacy analysis set	Included all randomised patients who had a baseline and ≥1 post-baseline evaluation of vertebral fracture at or before the timepoint of consideration. Patients were analysed according to their randomised treatment assignments. This was the primary analysis set for new, new or worsening, and multiple new or worsening vertebral fractures endpoints. Patients whose first post-baseline spinal radiograph showed no fracture on vertebra, but who had the same vertebrae at baseline were also included as it could be inferred that their baseline scores would have also reported no fracture, had they been available.
Safety analysis set	Patients who received ≥1 active dose of investigational product in the 12-month double-blind study period were included in this study set. Safety data analysis for the double-blind study period, primary analysis period, and overall study period used this safety analysis set.
Based on CS, Table 8, pages 34-35.[1] CS = company submission; MOF = major osteoporotic fracture	

A summary of the statistical tests that were used during primary analysis of ARCH, and the methods by which missing data were managed, is presented in Table 3.7. For new vertebral fractures through month 12 or month 24, and clinical and non-vertebral fractures through month 12, month 24 and to primary analysis, subgroup analyses were conducted for age, presence or absence of severe vertebral fracture at baseline, number of prevalent fractures at baseline, race, geographical region, Central/Latin

America and all regions excluding Central/Latin America, baseline lumbar spine BMD T-score, baseline total hip or femoral neck BMD T-score, Fracture Risk Assessment tool (FRAX) score and history of non-vertebral fracture at age ≥55 years. For change from baseline in BMD, subgroup analyses were conducted at month 12 and month 24 for age, geographical region, baseline BMD T-score at the lumbar spine and baseline BMD T-score at the total hip.

Table 3.7: Statistical tests for the primary analysis of ARCH

Trial number (acronym)	NCT01631214 (ARCH)
Hypothesis objective	Statistical hypothesis: 12 months treatment with romosozumab followed by alendronate is effective in reducing the incidence of a clinical fracture and new vertebral fracture in postmenopausal women with osteoporosis, compared to treatment with alendronate alone.
Statistical tests	Kaplan Meier estimates were used to summarise the cumulative incidence of fracture and a Cox proportional-hazards model stratified for age and prevalent vertebral fracture was used as a basis to assess treatment comparisons. A logistic regression model based on the primary efficacy analysis set for vertebral fractures was used to compare patient incidence of new vertebral fractures up to month 24. Adjusted odds ratio and the corresponding 95% CI were also given. To demonstrate the robustness of the primary analytical model results, additional supportive analysis was conducted including: per protocol analyses and time-to-event analysis based on full analysis set. The statistical significance for the primary and selected key secondary endpoints were controlled using sequential testing procedure to maintain the overall significance level for the study at 0.05. If both the primary endpoints were significant at the 0.05 level (2-sided), each of the following secondary DXA BMD endpoints were tested hierarchically at 0.05 (2-sided). With this procedure, formal inferential testing was performed for a step only when statistical significance was declared for all endpoints tested in previous steps. If the testing sequence stopped, the remaining endpoints in the testing sequence were not formally tested for statistical significance and the corresponding p-values were considered descriptive. The p-values for the analyses of other secondary, exploratory, and sub-study endpoints were nominal without adjusting for multiplicity. All p-values were 2-sided.
Data management, patient withdrawals	For BMD, missing data was dealt with by using LOCF. Patients who had missing data for a scheduled visit were not included in the safety data collections for that time point (no imputation). *Post hoc* analysis of vertebral fractures using a multiple-imputation method was performed for all randomly assigned patients. Observed data (excluding any imputed values) was reported through to 36 months including BMD scores at month 36.

Based on CS, Table 9, page 36.[1]
BMD = bone mineral density; CI = confidence interval; CS = company submission; DXA = dual-energy X-ray absorptiometry; LOCF = last-observation-carried-forward

ERG comment: The ERG has no particular concerns about the statistical analysis of the ARCH trial.

3.2.3 Baseline characteristics of the ARCH trial

In the ARCH trial, nearly all patients had experienced an osteoporotic fracture prior to the trial (99.1% in alendronate arm vs. 98.8% in romosozumab arm). Of the participants that were randomised to the

alendronate or romosozumab arms, a similar number had suffered non-vertebral fractures (13.4% vs. 13.2%) or vertebral fractures (25.2% vs. 27.7%), respectively, in the two years before enrolment. Participants had a mean age of approximately 74 years.[3] Baseline characteristics were comparable across both treatment groups. Key baseline demographics and clinical characteristics for the patients included in the full analysis set in ARCH are presented in Table 3.8.

Table 3.8: Baseline characteristics in the full analysis set in the ARCH trial

Characteristic	Alendronate (N=2,047)	Romosozumab (N=2,046)
Mean age, years (SD)	74.2 (7.5)	74.4 (7.5)
Age ≥75 years, no. (%)	1,071 (52.3)	1,073 (52.4)
Ethnic group, no. (%)		
Hispanic	662 (32.3)	631 (30.8)
Non-Hispanic	1,385 (67.7)	1,415 (69.2)
Geographical region, no. (%)		
Central or Eastern Europe or Middle East	798 (39.0)	835 (40.8)
Latin America	727 (35.5)	674 (32.9)
Western Europe, Australia, or New Zealand	264 (12.9)	269 (13.1)
Asia-Pacific or South Africa	216 (10.6)	213 (10.4)
North America	42 (2.1)	55 (2.7)
Mean BMI (SD)	25.36 (4.42)	25.46 (4.41)
Mean BMD T-score (SD)		
Lumbar spine	–2.99 (1.24)	–2.94 (1.25)
Total hip	–2.81 (0.67)	–2.78 (0.68)
Femoral neck	–2.90 (0.50)	–2.89 (0.49)
Previous osteoporotic fracture at ≥45 years of age, no. (%)	2,029 (99.1)	2,022 (98.8)
Prevalent vertebral fracture, no. (%)	1,964 (95.9)	1,969 (96.2)
Grade of most severe vertebral fracture[a]		
Mild	73 (3.6)	68 (3.3)
Moderate	570 (27.8)	532 (26.0)
Severe	1,321 (64.5)	1,369 (66.9)
Previous non-vertebral fracture at ≥45 years of age, no. (%)	770 (37.6)	767 (37.5)
Previous hip fracture, no. (%)[b]	179 (8.7)	175 (8.6)
Mean FRAX MOF risk (SD)	20.0 (10.1)	20.2 (10.2)
Median serum β-CTX, ng/l (IQR)[c]	230.0 (137.0–388.0)	276.0 (166.0–407.0)
Medium serum P1NP, µg/l (IQR)[c]	44.7 (32.7–64.4)	50.6 (37.5–64.7)
Median 25-hydroxyvitamin D, ng/ml (IQR)	27.6 (24.0–34.2)	28.4 (24.0–34.8)

Based on CS, Table 7, pages 33-34.[1]

[a] The grade of the most severe fracture was assessed with the use of the Genant grading scale.[6] [b] Previous hip fracture excludes pathologic or high-trauma hip fracture. [c] Data shown are for the 266 patients (128 in the

Characteristic	Alendronate (N=2,047)	Romosozumab (N=2,046)
alendronate group and 138 in the romosozumab group) who enrolled in the biomarker sub-study and who had measurements of bone-turnover markers both at baseline and at one or more visits after baseline. β-CTX = Beta-C-Terminal Telopeptide of Type 1 Collagen; BMD = bone mineral density; BMI = body mass index; CS = company submission; FRAX = Fracture Risk Assessment tool; IQR = interquartile range; MOF = major osteoporotic fracture; P1NP = Procollagen Type 1 N-Telopeptide; SD = standard deviation		

3.2.4 Risk of bias assessment of the ARCH trial

The RoB of the ARCH trial will be discussed in Section 3.3.4 of this report, together with the STRUCTURE and FRAME trials.

3.2.5 Efficacy results of the ARCH trial

The results from the ARCH trial presented in the CS describe those that were detailed in the ARCH clinical study report (CSR) and were determined using the standard methodology of last observation carried forward (LOCF) imputation for missing data, as pre-specified in the statistical analysis plan. The data more recently presented in the peer-reviewed New England Journal of Medicine publication regarding fractures and BMD were determined using a multiple imputation for the missing data,[3] as requested by the journal. As this does not reflect the original pre-specified analyses for the ARCH trial, the company did not include these results in their submission. The ERG asked the company to clarify whether there were any differences in estimates of effect between the two methods of imputation, and to describe how any differences between these analyses could affect the CE estimate (Clarification Letter, Question A13).[9] According to the company, the methodology used to derive the clinical effectiveness for vertebral fractures in the ARCH trial had no bearing on the results:

- Hazard ratio (HR) for new vertebral fractures at 12 months were 0.63, 95% CI 0.47 to 0.85 and 0.64, 95% CI 0.46 to 0.89) using multiple imputation and LOCF, respectively; and
- HR for new vertebral fractures at 24 months were 0.52 (0.40-0.66) and 0.50 (0.38-0.66) using multiple imputation and LOCF, respectively

Therefore, the results below will be based on the data presented in the CS.

In the ARCH trial, romosozumab/alendronate statistically significantly reduced the incidence of new vertebral fractures at month 24, meeting its primary endpoint. Patients in the romosozumab/alendronate arm had a 50% lower relative risk of vertebral fractures compared to patients on alendronate alone over 24 months (Table 3.9).[20] Additionally, a statistically significantly lower proportion of patients experienced a clinical fracture (non-vertebral fracture and clinical vertebral fracture) at the time of primary analysis in the romosozumab/alendronate group compared to alendronate alone, meeting the other primary endpoint.[20] Patients treated with romosozumab had a statistically significantly greater increase in BMD from baseline compared to alendronate (adjusted P<0.001), which was maintained until month 36 (Table 3.9).[20]

Table 3.9: Summary of clinical effectiveness results from ARCH

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
Primary outcomes			
Incidence of new vertebral	147/1834 (8.0%)	74/1825 (4.1%)	RR= 0.50 [redacted] (0.38, 0.66)

	Alendronate (N=2,047)	**Romosozumab (N=2,046)**	**Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)**
fracture at 24 months			
Incidence of clinical fracture at time of primary analysis (median 33 months)	266/2047 (13.0%)	198/2046 (9.7%)	HR= 0.73 [redacted] (0.61, 0.88)
Key secondary end points			
Incidence of non-vertebral fracture at the time of the primary analysis	217/2047 (10.6)	178/2046 (8.7)	HR= 0.81 [redacted] (0.66, 0.99)
BMD Outcomes: N, LS Mean (SE) – Mean Difference (95% CI)			
BMD at the lumbar spine at 12 months	[redacted], 5.0 ([redacted])	1722, 13.7 ([redacted])	MD = 8.7 (8.31, 9.09)
BMD at the lumbar spine at 24 months	[redacted], 7.2 ([redacted])	1571, 15.3 ([redacted])	MD = 8.1 (7.58, 8.57)
BMD at the lumbar spine at 36 months	[redacted], 7.8 ([redacted])	1593, 15.2 ([redacted])	MD = 7.4 (6.84, 7.89)
BMD at the total hip at 12 months	[redacted], 2.8 ([redacted])	1781, 6.2 ([redacted])	MD = 3.3 (3.03, 3.60)
BMD at the total hip at 24 months	[redacted], 3.5 ([redacted])	1622, 7.2 ([redacted])	MD = 3.8 (3.42, 4.10)
BMD at the total hip at 36 months	[redacted], 3.5 ([redacted])	1653, 7.2 ([redacted])	MD = 3.7 (3.29, 4.02)
BMD at the femoral neck at 12 months	[redacted], 1.7 ([redacted])	1781, 4.9 ([redacted])	MD = 3.2 (2.90, 3.54)
BMD at the femoral neck at 24 months	[redacted], 2.3 ([redacted])	1622, 6.0 ([redacted])	MD = 3.8 (3.40, 4.14)
BMD at the femoral neck at 36 months	[redacted], 2.4 ([redacted])	1653, 6.0 ([redacted])	MD = 3.6 (3.18, 3.97)

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
Other secondary end points			
Incidence of new vertebral fracture at 12 months	85/1703 (5.0%)	55/1696 (3.2%)	RR = 0.64 [redacted]; (0.46, 0.89)
Incidence of clinical fracture at 12 months	110/2047 (5.4)	79/2046 (3.9)	HR = 0.72 [redacted]; (0.54, 0.96)
Incidence of clinical fracture at 24 months	[redacted]	[redacted]	[redacted]
Incidence of non-vertebral fractures at 12 months	95/2047 (4.6)	70/2046 (3.4)	HR = 0.74 [redacted]; (0.54, 1.01)
Incidence of non-vertebral fractures at 24 months	[redacted]	[redacted]	[redacted]
Incidence of clinical vertebral fracture at 12 months	18/2047 (0.9)	10/2046 (0.5)	HR = 0.56 [redacted]; (0.26, 1.22)
Incidence of clinical vertebral fracture at 24 months	44/2047 (2.1)	18/2046 (0.9)	HR = 0.41 [redacted]; (0.24, 0.71)
Incidence of hip fractures at 12 months	22/2047 (1.1)	14/2046 (0.7)	HR = 0.64 [redacted]; (0.33, 1.26)
Incidence of hip fractures at 24 months	[redacted]	[redacted]	[redacted]
Incidence of hip fractures at primary analysis	66/2047 (3.2)	41/2046 (2.0)	HR = 0.62 [redacted]; (0.42, 0.92)
Incidence of major nonvertebral fractures at 12 months	88/2047 (4.3)	59/2046 (2.9)	HR = 0.67 [redacted]; (0.48, 0.94)
Incidence of major	196/2047 (9.6)	146/2046 (7.1)	HR = 0.73 [redacted]; (0.59, 0.90)

	Alendronate (N=2,047)	Romosozumab (N=2,046)	Risk ratio[a] (Point estimate (SE)[b]; (95% CI)) Hazard ratio[c] (SE) (95% CI)
nonvertebral fractures at primary analysis			
Incidence of major osteoporotic fractures at 12 months	85/2047 (4.2)	61/2046 (3.0)	HR = 0.72 [redacted]; (0.52, 1.01)
Incidence of major osteoporotic fractures at primary analysis	209/2047 (10.2)	146/2046 (7.1)	HR = 0.68 [redacted]; (0.55, 0.84)
Incidence of all osteoporotic fractures at 12 months	189/2047 (9.2)	134/2046 (6.5)	HR = 0.71 [redacted]; (0.57, 0.88)
Incidence of all osteoporotic fractures at primary analysis	392/2047 (19.1)	266/2046 (13.0)	HR = 0.65 [redacted]; (0.56, 0.76)

Based on CS, Section B.2.6, pages 38-43; CSR, Section 10.[1, 20]

[a] Values < 1 for RR favour romosozumab; based on the Mantel-Haenszel method adjusted for age strata, baseline total hip BMD T-score (≤-2.5, >-2.5), and presence of severe vertebral fracture at baseline; [b] SE represents the standard error of log (risk ratio); c Hazard ratio < 1 favours romosozumab; The HR estimate is based on Cox proportional hazards model adjusting for age strata, baseline total hip BMD T-score, and presence of severe vertebral fracture at baseline.

BMD = bone mineral density; CI = confidence interval; CS = company submission; HR = hazard ratio; MD = mean difference; RR = risk ratio; SE = standard error

As shown in Figure 3.2 there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first clinical fracture by month 12. At the time of primary analysis, patients treated with romosozumab/alendronate had a lower cumulative incidence of clinical fracture (9.7%) compared to the alendronate/alendronate group (13.0%) (nominal and adjusted P<0.001). This equated to a 27% lower relative risk of clinical fracture in the romosozumab/alendronate group than alendronate alone, meeting the co-primary endpoint for the ARCH trial.

ERG comment: Although the curves diverge from months zero to 42, they seem to converge again by month 48. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Longer term follow-up is needed to see whether the effects are maintained over time.

Figure 3.2: Kaplan-Meier curves for time to first clinical fracture

Based on CS, Figure 7, page 40.[1]
Footnote: Risks presented are based on a LOCF method for patients with missing fracture status. For Kaplan-Meier curves in the time-to-event analysis, data from patients who withdrew or reached the end of the reporting period without having a fracture were carried forward from the last observation time.
CS = company submission; LOCF = last observation carried forward; N = number of patients randomised; n = number of patients at risk for event at time point of interest

Similarly, patients treated with romosozumab showed a visible separation in time to non-vertebral fracture at month 12 compared to alendronate-treated patients, which was maintained for the duration of the study (Figure 3.3).[3]

ERG comment: Similar as in Figure 3.2, the curves in Figure 3.3 diverge from months 0 to 42 and seem to converge again by month 48. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Longer term follow-up is needed to see whether the effects are maintained over time.

Figure 3.3: Kaplan-Meier curves for time to first non-vertebral fracture

Based on CS, Figure 9, page 41.[1]
CS = company submission; N = number of subjects randomised; n = number of subjects at risk for event at time point of interest

ERG comment: Overall, results of the ARCH trial are favourable for romosozumab. Both primary outcomes (the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis) are met and most fracture results significantly favour romosozumab over alendronate. In addition, all BMD outcomes significantly favour romosozumab over alendronate. However, the graphs for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3), seem to indicate that the effectiveness of romosozumab over alendronate becomes less after 42 months; longer term follow-up is needed to see whether the effects are maintained over time.

3.2.5.1 Health-related quality of life

[redacted] in health-related quality of life (HRQoL) were observed between treatment groups in the ARCH trial.[1] According to the company, "this was to be expected because the HRQoL data were collected at predetermined, discrete time points irrespective of fracture occurrence during the trial and always related to one of the investigated treatments. Therefore, the trial data do not provide HRQoL values sensitive to decrease in HRQoL after a fracture, and are hence expected to underestimate the potential HRQoL gain with treatment".[1] The company also points out that it is *"important to note that the short nature of the trials meant that the analytical power for capturing HRQoL outcomes was limited"*.[1]

[redacted] By preventing fragility fractures, romosozumab (and alendronate) are expected to prevent future HRQoL decrements resulting from a fracture, according to the company.

3.2.6 Adverse events

3.2.6.1 Adverse events in the ARCH trial

The incidences of AEs and serious adverse events (SAEs) were similar overall in the ARCH trial between the two treatment groups during the 12-month double-blind period, and cumulative incidences were similar between the two groups during the primary analysis period (Table 3.10). In the first 12 months, injection-site reactions (mostly mild in severity) were reported in more patients receiving romosozumab (90 of 2,040 patients (4.4%)) than in those receiving alendronate (53 of 2,014 patients (2.6%)).

However, more people in the romosozumab group experienced adjudicated serious CV AEs during the double-blind period, with 50 patients (2.5%) in the romosozumab group and 38 (1.9%) in the alendronate group reporting these events (odds ratio, 1.31; 95% confidence interval (CI) 0.85 to 2.00). A total of 16 patients (0.8%) in the romosozumab group and 6 (0.3%) in the alendronate group reported cardiac ischemic events (odds ratio, 2.65; 95% CI, 1.03 to 6.77), and 16 patients (0.8%) in the romosozumab group and seven (0.3%) in the alendronate group reported cerebrovascular events (odds ratio, 2.27; 95% CI, 0.93 to 5.22) (Table 3.10).

Table 3.10: Adverse events in the ARCH trial

Event	**Month 12: Double-Blind Period**		**Primary Analysis: Double-Blind and Open-Label Period***	
	Alendronate (N = 2,014)	**Romosozumab (N = 2,040)**	**Alendronate to Alendronate (N = 2,014)**	**Romosozumab to Alendronate (N = 2,040)**
	number of patients (percent)			
Adverse event during treatment	1,584 (78.6)	1,544 (75.7)	1,784 (88.6)	1,766 (86.6)
Back pain†	228 (11.3)	186 (9.1)	393 (19.5)	329 (16.1)
Nasopharyngitis†	218 (10.8)	213 (10.4)	373 (18.5)	363 (17.8)
Serious adverse event	278 (13.8)	262 (12.8)	605 (30.0)	586 (28.7)
Adjudicated serious cardiovascular event‡	38 (1.9)	50 (2.5)	122 (6.1)	133 (6.5)
Cardiac ischemic event	6 (0.3)	16 (0.8)	20 (1.0)	30 (1.5)
Cerebrovascular event	7 (0.3)	16 (0.8)	27 (1.3)	45 (2.2)
Heart failure	8 (0.4)	4 (0.2)	23 (1.1)	12 (0.6)
Death	12 (0.6)	17 (0.8)	55 (2.7)	58 (2.8)
Noncoronary revascularisation	5 (0.2)	3 (0.1)	10 (0.5)	6 (0.3)
Peripheral vascular ischemic event not requiring revascularisation	2 (<0.1)	0	5 (0.2)	2 (<0.1)
Death	21 (1.0)§	30 (1.5)	90 (4.5)§	90 (4.4)

Event	Month 12: Double-Blind Period		Primary Analysis: Double-Blind and Open-Label Period*	
	Alendronate (N = 2,014)	Romosozumab (N = 2,040)	Alendronate to Alendronate (N = 2,014)	Romosozumab to Alendronate (N = 2,040)
	number of patients (percent)			
Event leading to discontinuation of trial regimen	64 (3.2)	70 (3.4)	146 (7.2)	133 (6.5)
Event leading to discontinuation of trial participation	27 (1.3)	30 (1.5)	43 (2.1)	47 (2.3)
Event of interest¶				
Osteoarthritis‖	146 (7.2)	138 (6.8)	268 (13.3)	247 (12.1)
Hypersensitivity	118 (5.9)	122 (6.0)	185 (9.2)	205 (10.0)
Injection-site reaction**	53 (2.6)	90 (4.4)	53 (2.6)	90 (4.4)
Cancer	28 (1.4)	31 (1.5)	85 (4.2)	84 (4.1)
Hyperostosis††	12 (0.6)	2 (<0.1)	27 (1.3)	23 (1.1)
Hypocalcaemia	1 (<0.1)	1 (<0.1)	1 (<0.1)	4 (0.2)
Atypical femoral fracture‡	0	0	4 (0.2)	2 (<0.1)
Osteonecrosis of the jaw‡	0	0	1 (<0.1)	1 (<0.1)

Based on Saag et al. 2017.[3]

* Incidence rates at the time of the primary analysis were cumulative and included all events in the double-blind and open-label period (to February 27, 2017) in patients who received at least one dose of open-label alendronate; † Shown are events that occurred in 10% or more of the patients in either group during the double-blind period; ‡ Serious cardiovascular adverse events were adjudicated by the Duke Clinical Research Institute, and potential cases of osteonecrosis of the jaw and atypical femoral fracture were adjudicated by independent committees. Cardiovascular deaths include fatal events that were adjudicated as being cardiovascular-related or undetermined (and, therefore, possibly cardiovascular-related); § One patient had a non–treatment-related serious adverse event of pneumonia that was incorrectly flagged as death in the primary analysis snapshot and was not included in the analysis of fatal events; ¶ Events of interest were those that were identified by prespecified Medical Dictionary for Regulatory Activities search strategies; ‖ Prespecified events that were reported under osteoarthritis were osteoarthritis, spinal osteoarthritis, exostosis, arthritis, polyarthritis, arthropathy, monoarthritis, and interspinous osteoarthritis; ** The most frequent adverse events of injection-site reactions (occurring in >0.1% of the patients) in the romosozumab group during the double-blind period included injection-site pain (in 1.6% of the patients), erythema (1.3%), pruritus (0.8%), haemorrhage (0.5%), rash (0.4%), and swelling (0.3%); †† Prespecified events reported under hyperostosis were exostosis (mostly reported as heel spurs), lumbar spinal stenosis, spinal column stenosis, cervical spinal stenosis, enostosis, extra skeletal ossification, and vertebral foraminal stenosis.

3.2.6.2 Pooled adverse events from seven romosozumab studies

The safety and tolerability of romosozumab was evaluated in a programme including seven clinical trials, exposing more than 7,500 patients to romosozumab. The safety data presented in this section is a pooled analysis of the studies listed in Table 3.11, which includes the BRIDGE trial in men.

Table 3.11: Overview of studies included in the pooled safety analysis

Study	Design	Number of patients included in safety set
FRAME	Multicentre, international, randomised, double-blind, placebo-controlled, parallel-group, Phase III	Safety analysis set (n=7,157)
ARCH	Multicentre, international, randomised, double-blind, active-controlled, Phase III	Safety analysis set (n=4,054)
NCT00896532	Dose-ranging, randomised, placebo- and active controlled in women with low BMD	Safety analysis set (n=410)
NCT01992159	Dose-ranging, placebo-controlled in Japanese postmenopausal women with osteoporosis	Safety analysis set (n=252)
STRUCTURE	Multicentre, international, randomised, open-label, active-controlled, parallel-group, Phase III	Safety analysis set (n=432)
NCT02016716	Placebo-controlled, noninferiority study of romosozumab 70 vs. 90 mg/mL in postmenopausal women with osteoporosis	Safety analysis set (n=294)
BRIDGE	Multicentre, international, randomised, double-blind, placebo-controlled, Phase III	Safety analysis set (n=244) Included the male osteoporosis population

Across the pooled safety analysis set, which included the studies outlined in Table 3.11, the incidence of treatment-emergent adverse events (TEAEs) was similar in patients treated with romosozumab compared to the control group (Tables 3.12 and 3.13); the control included patients treated with placebo, alendronate and teriparatide across the clinical trial programme; exposure-adjusted incidence rate per 100 patient years: [redacted] events per 100 years (romosozumab) vs. [redacted] events per 100 years (control). Treatment related SAEs leading to discontinuation of study drug were also comparable (Table 3.12; exposure-adjusted incidence rate per 100 patient years of [redacted] in both the control group and romosozumab 210 mg QM group).

In the pooled studies, [redacted]% of patients treated with 210 mg QM romosozumab reported a serious TEAE, compared to [redacted]% of patients in the control group (Table 3.12). The most common serious TEAE reported was pneumonia ([redacted]% romosozumab 210 mg QM-treated patients vs. [redacted]% control-treated patients).

Table 3.12: Summary of exposure-adjusted incidence rate of treatment emergent adverse events (osteoporosis safety analysis set)

	All Studies (Including ARCH)		
	Control[a] (N=[redacted]) n (r)	**Romosozumab 210 mg QM[b] (N=[redacted]) n (r)**	**Romosozumab Total[c] (N=[redacted]) n (r)**
All treatment-emergent adverse events			
All TEAEs	[redacted]	[redacted]	[redacted]
Serious AEs	[redacted]	[redacted]	[redacted]
Leading to discontinuation of investigational product	[redacted]	[redacted]	[redacted]
Fatal AEs*	[redacted]	[redacted]	[redacted]

	All Studies (Including ARCH)		
	Control[a] (N=[redacted]) n (r)	Romosozumab 210 mg QM[b] (N=[redacted]) n (r)	Romosozumab Total[c] (N=[redacted]) n (r)
Treatment-related treatment-emergent adverse events[d]			
Treatment-related TEAEs	[redacted]	[redacted]	[redacted]
Serious AEs	[redacted]	[redacted]	[redacted]
Leading to discontinuation of investigational product	[redacted]	[redacted]	[redacted]
Fatal AEs	[redacted]	[redacted]	[redacted]

Based on CS, Table 14, page 55.[1]

* Alendronate-treated subject 14248015041 had a fatal non-treatment-related serious AE of pneumonia that had an incorrect death flag in the primary analysis snapshot and was not included in the exposure-adjusted incidence rate of fatal events; a Includes placebo from Studies FRAME (12 months), NCT00896532 (24 months), NCT01992159 (12 months), BRIDGE (12 months), and NCT02016716 (6 months), alendronate from Studies NCT00896532 (12 months) and ARCH (12 months), and teriparatide from studies NCT00896532 (12 months), and STRUCTURE (12 months); b Includes Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months), NCT01992159 (12 months), 20110142 (12 months), BRIDGE (12 months), and NCT02016716 (6 months); c Includes romosozumab QM and Q3M from Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months, all data), NCT01992159 (12 months), ARCH (12 months), BRIDGE (12 months), and NCT02016716 (6 months); d Includes only events for which the investigator indicated there was a reasonable possibility they may have been caused by investigational product.
AE = adverse event; CS = company submission; QM = every month; QW = every week; r = exposure-adjusted incidence rate per 100 subject-years; TEAE = treatment-emergent adverse event

Table 3.13: Exposure-adjusted incidence rate of most frequent (≥5.0 per 100 subject-years in total romosozumab or integrated control groups) adverse events by preferred term (osteoporosis safety analysis set)

Preferred term*	All Studies (Including ARCH)		
	Control[a] (N=[redacted]) n (r)	Romosozumab 210 mg QM[b] (N=[redacted]) n (r)	Romosozumab Total[c] (N=[redacted]) n (r)
Number of patients reporting treatment-emergent AEs	[redacted]	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]
Pain in extremity	[redacted]	[redacted]	[redacted]
Fall	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]
Osteoarthritis	[redacted]	[redacted]	[redacted]

Upper respiratory tract infection	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]
Viral upper respiratory tract infection	[redacted]	[redacted]	[redacted]

Source: CS, Table 15, page 56.[1]

* Preferred terms are sorted by descending order of the exposure-adjusted incidence rate in the total romosozumab group and control group and coded using Medical Dictionary for Regulatory Activities version 19.1; a Includes placebo from Studies FRAME (12 months), NCT00896532 (24 months), NCT01992159 (12 months), BRIDGE (12 months), and NCT02016716 (6 months), alendronate from Studies NCT00896532 (12 months) and ARCH (12 months) and teriparatide from Studies NCT00896532 (12 months), and STRUCTURE (12 months); b Includes Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months), NCT01992159 (12 months), ARCH (12 months), BRIDGE (12 months), and NCT02016716 (6 months); c Includes romosozumab QM and Q3M from Studies FRAME (12 months), NCT00896532 (24 months), STRUCTURE (12 months, all data), NCT01992159 (12 months), 20110142 (12 months), BRIDGE (12 months), and NCT02016716 (6 months).

AE = adverse event; CS = company submission; QM = every month; QW = every week; r = exposure-adjusted incidence rate per 100 subject-years; TEAE = treatment-emergent adverse event

3.2.7 Included studies: Supporting evidence

According to the company, the clinical effectiveness evidence for romosozumab in severe osteoporosis is provided from three phase III clinical trials: ARCH, FRAME and STRUCTURE. A fourth study, BRIDGE, considered use in men, which is not part of the marketing authorisation for romosozumab; as such, no clinical effectiveness results are presented from BRIDGE in the CS.[1] However, some data from BRIDGE are introduced in the pooled safety analysis (see Section 3.2.6 of this report).

The ARCH trial has been discussed in the sections above. Neither the FRAME nor STRUCTURE trials studied a patient population aligned to where romosozumab is expected to be used in NHS clinical practice. In addition, STRUCTURE was also not designed to evaluate fracture outcomes.[17, 18] Therefore, the FRAME and STRUCTURE trials will only be minimally discussed in this section of the ERG report.

Table 3.14: Supporting evidence

Study	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
Study design	International, multicentre, randomised, double-blind, placebo-controlled, parallel-group, Phase III.	International, multicentre, randomised, open-label, active-controlled, parallel-group, phase III.
Population	• Postmenopausal women with osteoporosis • Aged 55–90 years	• Postmenopausal women with osteoporosis transitioning from bisphosphonate therapy • Aged 55–90 years • Prior fragility fracture
Intervention(s)	Romosozumab (210 mg) QM SC for 12 months followed by open-label denosumab (60 mg) SC Q6M for 24 months (until study end).	Romosozumab (210 mg) QM SC for 12 months.
Comparator(s)	Placebo QM SC for 12 months followed by open-label denosumab (60 mg) Q6M SC for 24 months (until study end).	Daily SC teriparatide (20 µg) for 12 months.
Reported outcomes relevant to the decision problem	• Incidence of a new vertebral fracture • Cumulative incidence of non-vertebral fracture, major non-vertebral fracture, clinical fracture,	• Percent change from baseline in BMD at LS, TH, and FN • Finite element analysis of the hip[a]

Study	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
	hip fracture, new or worsening vertebral fracture, MOF and multiple new or worsening vertebral fractures • Percent change from baseline in BMD at LS, TH, and FN • EQ-5D-5L, OPAQ-SV, LAD, and BPI worst pain • AEs	• AEs
Based on CS, Table 4, page 29.[1] AE = adverse event; BMD = bone mineral density; BPI = Brief Pain Inventory; CS = company submission; EQ-5D-5L = EuroQoL-5 Dimensions-5 Levels Health Survey; FN = femoral neck; LAD = limited activity days; LS = lumbar spine; MOF = major osteoporotic fracture; OPAQ-SV = Osteoporosis Assessment Questionnaire Short Version; SC = subcutaneous; TH = total hip; Q6M = once every six months; QM = once monthly; TH = total hip		

3.2.7.1 The FRAME Study

The FRAME study demonstrated statistically significant reductions in new vertebral fractures for romosozumab compared with placebo at 12 months follow-up (relative risk reduction (RRR): 73%; absolute risk reduction (ARR): 1.30%; adjusted P<0.001). Similarly, patients in the romosozumab/denosumab arm showed a statistically significant 75% reduction in RR of new vertebral fracture compared to the placebo/denosumab arm (ARR: 1.89%; incidence of new vertebral fracture: 0.6% vs. 2.5%; 95% CI: 60 to 84; adjusted P<0.001) at 24 months follow-up.[17] Romosozumab also reduced the risk of clinical fracture (non-vertebral and clinical vertebral fracture) by 36% compared with placebo at 12 months follow-up (adjusted and nominal P=0.008) and to 33% at 24 months follow-up (adjusted P=0.096, nominal P=0.002).[17]

3.2.7.2 The STRUCTURE Study

The STRUCTURE study provides BMD and estimated bone strength data comparing romosozumab and teriparatide in a population with severe osteoporosis and who received an oral bisphosphonate before transitioning to the bone-forming agent. In the STRUCTURE study, the mean percentage change from baseline up to month 12 in BMD at the total hip was 3.2% higher (95% CI: 2.7 to 3.8; adjusted P<0.0001) in the romosozumab group (2.6%, 95% CI: 2.2 to 3.0) compared to teriparatide (–0.6%, 95% CI –1.0 to –0.2).[18]

3.2.8 Ongoing studies

Three ongoing Post-Authorization Safety Studies (PASS) in the European Union (EU), one in the United States of America (USA) and another in South Korea are proposed to evaluate adherence to the risk minimisation measures in the romosozumab SmPC; to evaluate potential differences in serious cardiovascular AEs between romosozumab and currently-available therapies in real-world conditions; and to evaluate potential difference in serious infections between romosozumab and currently-available therapies in real-world conditions, respectively. The studies will use a multi-database approach with routinely collected data and are expected to last for a period of six years. The company is also aiming to conduct a study to assess the efficacy and safety of romosozumab in Chinese patients.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company conducted NMAs to compare the efficacy of romosozumab and romosozumab/alendronate and other bisphosphonates (alendronate, risedronate, ibandronate, zoledronate), teriparatide, denosumab and raloxifene. The ARCH, FRAME and STRUCTURE studies

contributed information for the direct comparisons between romosozumab and romosozumab/alendronate with alendronate, teriparatide and placebo. Other studies comparing comparator treatments with placebo and other comparator treatments were found using the systematic review described in Section 3.1.

Five distinct outcomes were considered in the NMAs: 1) new vertebral fractures at 12, 24 and 36 months, 2) non-vertebral fractures at 12, 24 and 36 months, 3) hip fractures at 12, 24 and 36 months, lumbar spine BMD at each study's latest timepoint, 4) total hip BMD at each study's latest timepoint, and 5) femoral neck BMD at each study's latest timepoint. For fracture outcomes, results were available both for the ITT population (base-case) and the EU label population; in this report, we will focus on the ITT population results only.

The inclusion and exclusion criteria of the NMAs are shown in Table 3.15.

Table 3.15: The inclusion and exclusion criteria of the NMAs

PICOS criteria	Inclusion criteria	Exclusion criteria
Population	Postmenopausal women with osteoporosis	Did not report on the population of interest.
Interventions or Comparators	Studies comparing at least two interventions of interest (plus background therapy, defined as calcium supplements and/or vitamin D): • Placebo • Romosozumab (210 mg SC QM) • Romosozumab & Alendronate (ROMO & ALN) - 210 mg SC QM & 70 mg QW • Raloxifene (60 mg oral QD) • Alendronate (10 mg oral QD or 70mg oral QW) • Risedronate (5 mg oral QD or 35mg oral QW) • Zoledronate (5 mg IV yearly) • Denosumab (60 mg SC twice yearly) • Teriparatide (20 µg SC QD)* • Abaloparatide (80 mg SC QD) • Ibandronate* (150 mg oral QM)	Did not compare at least two relevant interventions.

Outcomes	Studies reporting appropriate data for one of the following outcomes.
	• Fracture outcomes at 12, 24 and 36 months:
	  • New vertebral fracture
	  • Nonvertebral fracture
	  • Hip fracture
	• BMD outcomes (percentage change at the latest time point available from each trial):
	  • Femoral neck
	  • Lumbar spine
	  • Total hip
		Did not report any relevant outcomes or did not report appropriate data (e.g., RR but no 95% CrI, SD or SE).

Based on CS, Table 22 of Appendix D.[8]
* Ibandronate was included only in the BMD outcomes. ** One trial (i.e., Hadji et al. 2012) reported on a teriparatide dose of 20 µg SC QW.
ALN = alendronate; BMD = bone mineral density; CrI = credible interval; CS = company submission; IV = intravenous; PICOS = population, intervention, comparator, outcome, study design, QD = once daily; QM = once monthly; QW = once weekly; ROMO = romosozumab; RR = relative risk; SC = subcutaneous; SD = standard deviation; SE = standard error

3.3.1 Details of the trials included in the NMAs

Different studies were included in each network for each outcome and timepoint depending on the data available, though there were similarities across networks. Networks for all fracture outcomes at 12 months used ARCH[21] for the direct comparison between romosozumab and alendronate and FRAME[22] for the direct comparison between romosozumab and placebo. Networks for fractures at 24 and 36 months used ARCH for the direct comparison between romosozumab/alendronate and alendronate. Therefore, for fracture outcomes, only indirect evidence is available for comparisons of romosozumab and romosozumab/alendronate with comparator treatments other than alendronate and placebo (at 12 months). Most studies in the NMAs for fracture outcomes compared a comparator treatment with placebo, meaning consistency cannot be assessed for most comparisons. This is because inconsistency is assessed by comparing direct and indirect comparisons of treatments, which requires a loop in a network (the simplest being a triangle, with direct evidence linking three treatments). As the vast majority of comparisons between romosozumab and comparator treatments in all NMAs only have indirect evidence, inconsistency cannot be assessed.

Networks for all BMD outcomes used ARCH[21] for the direct comparison between romosozumab and alendronate, FRAME[22] for the direct comparison between romosozumab and placebo and STRUCTURE[18] for the direct comparison between romosozumab and teriparatide. Therefore, for BMD outcomes, only indirect evidence is available for comparisons of romosozumab and romosozumab/alendronate with comparator treatments other than alendronate, teriparatide and placebo. There were more comparisons with comparator treatments other than placebo in the BMD NMAs, meaning both direct and indirect evidence if available, and so consistency could be checked for more comparisons.

Tables 3.16 and 3.17 show a list of the comparator treatments and timepoints available for each outcome for studies included in at least one network for fracture and BMD outcomes, respectively.

Table 3.16: Studies included in the NMAs of fracture outcomes

Trial/Study	Intervention	Comparator 1	Comparator 2	Included in ITT analysis	Included in EU label-matched analysis	New vertebral timepoints	Non-vertebral timepoints	Hip timepoints
ACTIVE trial[23]	Abaloparatide	Placebo	Teriparatide	Yes	Yes	24	24	24
ARCH trial[21]	Romosozumab	Alendronate	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Bai et al. 2013[12]	Zoledronate	Placebo	NA	Yes	Yes	24		24
Chao et al. 2013[24]	Zoledronate	Placebo	NA	Yes	Yes	NA	12, 36	12, 36
Dursun et al. 2001[25]	Alendronate	Placebo	NA	Yes	Yes	12	NA	NA
EVA trial	Alendronate	Placebo	NA	Yes	Yes	NA	NA	12
FIT I + II trial[26]	Alendronate	Placebo	NA	Yes	Yes	NA	36	
FIT I trial[27]	Alendronate	Placebo	NA	Yes	Yes	24, 36	NA	24, 36
FOSIT trial[28]	Alendronate	Placebo	NA	Yes	Yes	NA	12, 24	NA
FRAME trial[22]	Romosozumab	Placebo	NA	Yes	Yes	12	12	12
FREEDOM trial[15]	Denosumab	Placebo	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Hadji et al. 2012[29]	Teriparatide	Risedronate	NA	Yes	Yes	24	24	24
HORIZON-PFT trial[13]	Zoledronate	Placebo	NA	Yes	Yes	12, 24, 36	12, 24, 36	12, 24, 36
Liberman et al. 1995[30]	Alendronate	Placebo	NA	Yes	Yes	36	NA	NA
Liu et al. 2004[31]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
Lufkin et al. 1998[32]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
MORE trial[33]	Raloxifene	Placebo	NA	Yes	Yes	12, 24, 36	NA	NA
Morii et al. 2003[14]	Raloxifene	Placebo	NA	Yes	Yes	12	NA	NA
Neer et al. 2001[34]	Teriparatide	Placebo	NA	Yes	Yes	24	12, 24	24
ROSE trial[35]	Alendronate	Zoledronate	NA	Yes	Yes	NA	NA	NA
RUTH trial[36]	Raloxifene	Placebo	NA	Yes	Yes	NA	12, 24, 36	12, 24, 36
Silverman et al. 2008 (93)[37]	Raloxifene	Placebo	NA	Yes	Yes	36	36	NA

Trial/Study	Intervention	Comparator 1	Comparator 2	Included in ITT analysis	Included in EU label-matched analysis	New vertebral timepoints	Non-vertebral timepoints	Hip timepoints
VERO trial[38]	Teriparatide	Risedronate	NA	Yes	Yes	12, 24	12, 24	24
VERT MN trial (EU analysis)[39]	Risedronate	Placebo	NA	Yes	Yes		12, 24, 36	36
VERT-MN trial (AUS+EU analysis)[39]	Risedronate	Placebo	NA	Yes	Yes	12, 24, 36	NA	NA
VERT-MN trial (NAm analysis)[40]	Risedronate	Placebo	NA	Yes	Yes	12, 36	36	36
ZONE trial[41]	Zoledronate	Placebo	NA	Yes	Yes	24	NA	NA

Based on Table 24 of Appendix D of the CS.[8]
*Patients switched to alendronate after 24 months. **Patients switched to denosumab after 12 months.
Dosing schedules: Placebo, romosozumab (210 mg SC QM), raloxifene (60 mg oral QD), alendronate (10 mg oral QD or 70 mg oral QW), risedronate (5 mg oral QD or 35 mg oral QW), zoledronate (5 mg IV yearly), denosumab (60 mg SC twice yearly), teriparatide (20 µg SC QD (QW for Hadji et al. 2012)), abaloparatide (80 mg SC QD).
AUS = Australia; CS = company submission; EU = European Union; ITT = intention-to-treatment; IV = intravenous; NA = not applicable; NAm = North America; NMA = network meta-analysis; QD = once daily; QM = once monthly; QW = once weekly; SC = subcutaneous

Table 3.17: Studies included in the NMAs of BMD outcomes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
ACTIVE trial[23]	Abaloparatide	Teriparatide	Placebo		Yes	Yes	Yes
Adami et al. 1995[42]	Alendronate	Placebo			Yes	Yes	Yes
Adami et al. 2008[43]	Raloxifene	Placebo			Yes	No	Yes
Aki et al. 2004[44]	Alendronate	Placebo			Yes	No	Yes
Amgen 20010223[45]	Denosumab	Alendronate	Placebo		Yes	Yes	No
ARCH[21]	Romosozumab	Alendronate			Yes	Yes	Yes
DATA[46]	Denosumab	Teriparatide			Yes	Yes	Yes
DECIDE[47]	Denosumab	Alendronate			Yes	Yes	Yes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
DEFEND[48]	Denosumab	Placebo			Yes	Yes	Yes
Dursun et al. 2001[25]	Alendronate	Placebo			Yes	No	Yes
EFFECT[49]	Raloxifene	Alendronate			Yes	No	No
EFFECT international[50]	Alendronate	Raloxifene			Yes	Yes	Yes
EUROFORS[51]	Teriparatide	Raloxifene	Placebo		Yes	Yes	Yes
FACT[52]	Alendronate	Risedronate			Yes	Yes	Yes
FACTS1[53]	Alendronate	Risedronate			Yes	Yes	Yes
Fogelman et al. 2000[54]	Risedronate	Placebo			Yes	No	Yes
FOSIT[28]	Alendronate	Placebo			Yes	Yes	Yes
FRAME[22]	Romosozumab	Placebo			Yes	Yes	Yes
Grey et al. 2010[55]	Zoledronate	Placebo			Yes	Yes	No
Hadji et al. 2012[29]	Risedronate	Teriparatide			Yes	Yes	Yes
HORIZON[13]	Zoledronate	Placebo			Yes	Yes	Yes
Johnell et al. 2002[56]	Raloxifene	Alendronate	Placebo		Yes	No	Yes
Liberman et al. 1995[30]	Alendronate	Placebo			Yes	No	Yes
McClung et al. 2009[57]	Ibandronate	Placebo			No	Yes	Yes
McClung et al. 2014[58]	Romosozumab	Teriparatide	Alendronate	Placebo	Yes	Yes	Yes
Miller et al. 2016[59]	Denosumab	Zoledronate			Yes	Yes	No
MOTION[60]	Ibandronate	Alendronate			Yes	Yes	No
NCT00132808[61]	Zoledronate	Placebo			Yes	Yes	Yes
NCT00353080[61]	Risedronate	Placebo			Yes	No	No
NCT00398606[62]	Alendronate	Placebo			Yes	No	Yes
Neer et al. 2001[34]	Teriparatide	Placebo			Yes	Yes	Yes

Studies	Intervention	Comparator			BMD		
					Lumbar spine	Total hip	Femoral neck
	Arm 1	Arm 2	Arm 3	Arm 4	Availability of data per BMD endpoint		
OCEAN[63]	Alendronate	Placebo			Yes	No	No
Recknor et al. 2013[64]	Denosumab	Ibandronate			Yes	Yes	Yes
Reid et al. 2011[65]	Zoledronate	Placebo			Yes	No	No
Roux et al. 2013[66]	Denosumab	Risedronate			Yes	Yes	Yes
Silverman et al. 2008[37]	Raloxifene	Placebo			Yes	Yes	No
SPIMOS[67]	Ibandronate	Placebo			Yes	Yes	No
STAND[68]	Denosumab	Alendronate			Yes	Yes	No
STRUCTURE[18]	Romosozumab	Teriparatide			Yes	Yes	Yes
Tan et al. 2016[69]	Zoledronate	Alendronate			Yes	Yes	Yes
Tucci et al. 1996[70]	Alendronate	Placebo			Yes	Yes	Yes
Um et al. 2017[71]	Raloxifene	Alendronate	Placebo		Yes	No	Yes

Source: Table 25 of Appendix D of the CS.[8]
BMD = bone mineral density; CS = company submission

3.3.2 Statistical analyses of the NMAs

For the fracture outcomes, RRs were used to estimate the relative effectiveness of all treatments, based on the number of participants in each treatment group in each study and the number of participants developing fractures by each timepoint. For BMD outcomes, mean differences (MDs) with 95% credible intervals (CrIs) were used to estimate the relative effectiveness of treatments. Some studies were missing data for the specified timepoints (12, 24 and 36 months), and were included if there were other informative timepoints, e.g. for new vertebral fractures, the ACTIVE study[23] had results at 18 months comparing abaloparatide and teriparatide, which was included in the 24-month NMA. Additionally, data from FRAME was only used at 12 months, as after 12 months, all patients in FRAME switched to denosumab.

The NMAs were conducted in a Bayesian framework: binary Bayesian models were used for fracture outcomes and shared parameter Bayesian models were used for BMD outcomes. Non-informative priors were used for all analyses. Both fixed and random effects models were presented for fracture outcomes, but only random effects models were presented for BMD outcomes due to high levels of heterogeneity observed in previous NMAs. All NMAs were run with 50,000 iterations after a burn-in of 30,000 iterations. An additional 50,000 iterations were run if the data were not sufficient converged after the initial 50,000 iterations, based on NMA diagnostic. All presented results converged.

Homogeneity was assessed using the I^2 statistic, using threshold values to indicate little (zero to 40%), moderate (30% to 60%), substantial (50% to 90%) and considerable (75% to 100%) heterogeneity. Consistency was assessed using the Bucher method, taking a P value of <0.05 as significant inconsistency, though no further action was taken in the presence of inconsistency. Baseline characteristics were compared to assess similarity of included studies, including mean age, the proportion of subjects with prevalent fracture, and mean BMD. Publication bias was not assessed.

Results were presented as tables comparing all comparator treatments, as ranks for all comparator treatments (the percentage chance of having the top, second, third rank etc.), and as forest plots showing the effectiveness of comparator treatments relative to romosozumab or romosozumab/alendronate.

3.3.3 Baseline characteristics of the trials in the NMAs

Table 3.18 details the intervention and comparator treatments for all trials included in any of the fracture NMAs, along with the outcomes and timepoints for which there were data.

Table 3.18: Trial details for all trials in any NMA of fracture outcomes

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
ACTIVE trial[23]	Abaloparatide	Placebo, Teriparatide		■			■			■	
ARCH trial[21]	Romosozumab	Alendronate	■	■	■	■	■	■	■	■	■
Bai et al. 2013[12]	Zoledronate	Placebo		■						■	
Chao et al. 2013[24]	Zoledronate	Placebo				■		■	■		■

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
Dursun et al. 2001[25]	Alendronate	Placebo	■								
EVA trial	Alendronate	Placebo							■		
FIT I + II trial[26]	Alendronate	Placebo						■			
FIT I trial[27]	Alendronate	Placebo		■	■					■	■
FOSIT trial[28]	Alendronate	Placebo				■	■				
FRAME trial[22]	Romosozumab	Placebo	■			■			■		
FREEDOM trial[15]	Denosumab	Placebo	■	■	■	■	■	■	■	■	■
Hadji et al. 2012[29]	Teriparatide	Risedronate		■			■			■	
HORIZON -PFT trial[13]	Zoledronate	Placebo	■	■	■	■	■	■	■	■	■
Liberman et al. 1995[30]	Alendronate	Placebo			■						
Liu et al. 2004[31]	Raloxifene	Placebo	■								
Lufkin et al. 1998[32]	Raloxifene	Placebo	■								
MORE trial[33]	Raloxifene	Placebo	■	■	■						
Morii et al. 2003[14]	Raloxifene	Placebo	■								
Neer et al. 2001[34]	Teriparatide	Placebo		■		■	■			■	
ROSE trial[35]	Alendronate	Zoledronate									
RUTH trial[36]	Raloxifene	Placebo				■	■	■	■	■	■
Silverman et al. 2008 (93)[37]	Raloxifene	Placebo			■			■			
VERO trial[38]	Teriparatide	Risedronate	■	■		■	■			■	
VERT MN trial (EU analysis)[39]	Risedronate	Placebo				■	■	■			■
VERT-MN trial	Risedronate	Placebo	■	■	■						

Trial/ Study	Intervention	Comparator	New vertebral timepoints (months)			Non-vertebral timepoints (months)			Hip timepoints (months)		
			12	24	36	12	24	36	12	24	36
(AUS+EU analysis)[39]											
VERT-MN trial (NAm analysis)[40]	Risedronate	Placebo									
ZONE trial[41]	Zoledronate	Placebo									

Based on Table 24 of Appendix D of the CS.[8]
CS = company submission

The company did not provide information for patient characteristics for included trials providing non-romosozumab evidence in any of the NMAs, though this information is crucial for determining whether there is a RoB in any individual comparison within an NMA. For NMAs to be unbiased, effect modifiers must be balanced across all included studies. This is particularly true if the treatment comparisons only include indirect evidence, as checking for inconsistency between direct and indirect evidence (for example, from unbalanced effect modifiers) is impossible.

The ERG has compiled a table showing the patient characteristics for all trials included in any NMA of fracture outcomes, Table 3.19. All data is taken from the original study reports reference by the company, and includes the inclusion/exclusion criteria, mean age, ethnicity and prevalence of vertebral fractures at baseline.

Table 3.19: Patient characteristics for all trials included in any NMA of fracture outcomes

Trial/study	Patient characteristics
ACTIVE trial[23]	Inclusion criteria: Postmenopausal women aged 49 to 86 years were eligible if they had BMD by dual energy x-ray absorptiometry T score of less than or equal to −2.5 and greater than −5.0 at the lumbar spine or femoral neck together with radiologic evidence of at least two mild vertebral fractures or at least one moderate vertebral fracture or history of a low-trauma fracture of the forearm, humerus, sacrum, pelvis, hip, femur, or tibia within the past 5 years. Women older than 65 years who met fracture criteria but had a T score of less than or equal to −2.0 and greater than −5.0 were eligible. Women older than 65 years were eligible without fracture criteria if either BMD T score was less than or equal to −3.0 and greater than −5.0. Eligibility required normal serum values for calcium, intact parathyroid hormone, phosphorus, and alkaline phosphatase and a 25-hydroxyvitamin D level of greater than 15 ng/mL (37.5 nmol/l (SI conversion, multiply by 2.496)). Exclusion criteria: Women were excluded if they had more than four mild, moderate, or any severe vertebral fractures (consistent with definitions described by Genant et al), fewer than two evaluable lumbar vertebrae, or if hip BMD was unevaluable. Participants were ineligible if they had evidence of metabolic bone disease or malabsorption or were taking any medications that would interfere with bone metabolism. Women were also excluded if they used bisphosphonates for more than 3 months in the past 5 years or denosumab within the past year. Women with a history of osteosarcoma were also excluded. Mean age: 69 years Ethnicity: White (80%); Asian (16%); Black or African American (3%); Other (1%) Prevalent vertebral fracture: 24%
ARCH trial[21]	Inclusion criteria: Ambulatory postmenopausal women 55 to 90 years of age who met at least one of the following criteria were eligible: a BMD T score of –2.5 or less at the total hip or femoral neck and either one or more moderate or severe vertebral fractures or two or more mild vertebral fractures; or a BMD T score of –2.0 or less at the total hip or femoral neck and either two or more moderate or severe vertebral fractures or a fracture of the proximal femur sustained 3 to 24 months before randomisation. Exclusion criteria: Patients with severe osteoporosis, an inability to take alendronate oral tablets or contraindications to alendronate, including a glomerular filtration rate below 35 ml per minute per 1.73 m^2 of body-surface area. Mean age: 74 years Ethnicity: Hispanic (32%); non-Hispanic (68%) Prevalent vertebral fracture: 96%
Bai et al. 2013[12]	Inclusion criteria: For inclusion in the study women with a primary diagnosis of osteoporosis had to be postmenopausal, have a BMD T-score <=2.5 at the femoral neck but no evidence of vertebral fractures, or a BMD T-score ≤1.5 with radiological diagnosis of two or more vertebral fractures. Exclusion criteria: (i) patients with secondary osteoporosis or other diseases known to affect bone metabolism; (ii) patients taking sodium fluoride, parathyroid hormone, anabolic steroids or growth hormone within six months of study entry, or systemic corticosteroids within

Trial/study	Patient characteristics
	12 months of study entry; (iii) patients with malignant, hepatic and renal diseases; and (iv) a serum calcium concentration of >11.0 mg/dl and untreated hypocalcaemia. Mean age: 57 years Ethnicity: Not stated (study conducted in one hospital in China) Prevalent vertebral fracture: 61%
Chao et al. 2013[24]	Inclusion criteria: Female patients diagnosed with osteoporosis. Exclusion criteria: Patients with secondary osteoporosis or other diseases which were known to affect bone metabolism were excluded. Patients taking anabolic steroids, sodium fluoride, and parathyroid or growth hormone within 6 months were also excluded. Patients who had malignant neoplasm, serum calcium more than 11.0 mg/dl, or untreated hypocalcaemia were also excluded. Mean age: 55 years Ethnicity: Not stated (study conducted in two hospitals in China) Prevalent vertebral fracture: 55%
Dursun et al. 2001[25]	Inclusion criteria: Postmenopausal women with a BMD of two SDs or more below the young adult mean at either the posteroanterior lumbar spine or the femoral neck. Exclusion criteria: Women with a documented history of drug or alcohol abuse, or with evidence from physical examination, laboratory tests or radiography of any bone metabolism disorder. Exclusion criteria also included active GI or liver disease, renal failure, renal calculi, treatment with specific therapy for osteoporosis, treatment with systemic corticosteroid therapy, malignancy, disorder of calcium metabolism and lumbar vertebrae abnormalities preventing the evaluation of BMD. Mean age: 61 years Ethnicity: Not stated (study conducted in one hospital in Turkey) Prevalent vertebral fracture: Not stated
FIT I + II trial[26]	Inclusion criteria: Women aged 55 to 80 years who had been post-menopausal for at least 2 years and had femoral neck BMD of 0.68 g/cm^2 or less. Exclusion criteria: Women with recent peptic ulcers or ulcers that required hospitalisation, dyspepsia requiring daily treatment, significant renal or hepatic dysfunction, medical problems that precluded three years of participation, severe malabsorption, blood pressure exceeding 210 mmHg systolic or 105 mmHg diastolic, MI within 6 months, unstable angina, hypothyroidism, hyperthyroidism, or hyperparathyroidism. Women taking oestrogen or calcitonin within the preceding six months or bisphosphonates or sodium fluoride (>1 mg/d) at any time were also excluded. Mean age: 68 years Ethnicity: Not stated (study conducted in 11 hospitals in the USA)

Trial/study	Patient characteristics
	Prevalent vertebral fracture: 0%
FIT I trial[27]	Inclusion criteria: Women aged 55 to 81 years who had been post-menopausal for at least 2 years and had femoral neck BMD of 0.68 g/cm^2 or less. Exclusion criteria: Women with recent peptic ulcers or ulcers that required hospitalisation, dyspepsia requiring daily treatment, significant renal or hepatic dysfunction, medical problems that precluded 3 years of participation, severe malabsorption, blood pressure exceeding 210 mmHg systolic or 105 mmHg diastolic, MI within 6 months, unstable angina, hypothyroidism, hyperthyroidism, or hyperparathyroidism. Women taking oestrogen or calcitonin within the preceding 6 months or bisphosphonates or sodium fluoride (>1 mg daily for 2 weeks or longer) at any time were also excluded. Mean age: 71 years Ethnicity: Not stated (study conducted in 11 hospitals in the USA) Prevalent vertebral fracture: 100%
FOSIT trial[28]	Inclusion criteria: Women eligible for study participation had been postmenopausal for at least 3 years, were not older than 85 years, and had BMD of the lumbar spine (L2–4) at least two standard deviations (SD) below the mean for mature, premenopausal women. Eligible patients were otherwise in good health and were between 20% below and 50% above ideal body weight as defined in the Metropolitan Life Insurance Company Height and Weight Table. Exclusion criteria: Excluded from participation were women with metabolic bone disease other than postmenopausal osteoporosis; disturbed parathyroid or thyroid function; major GI disease (for example, peptic ulcer or malabsorption) within the year before enrolment or use of a drug to inhibit gastric acid secretion for >2 weeks within 3 months of study entry; MI within the year prior to enrolment; uncontrolled hypertension or untreated angina; significantly impaired renal function (serum creatinine >150 mmol/l); or evidence of significant end organ disease. Also excluded were women who had received a bisphosphonate or fluoride (>8 mg/day) during the previous 6 months; oestrogen (except vaginal 43 times per week), ipriflavone or calcitonin during the previous 4 months; or any anabolic steroid, glucocorticoid or progestin for >2 weeks within the previous 6 months. Participants could not be receiving any medications that might alter bone or mineral metabolism, including vitamin A in excess of 10,000 U/day, vitamin D in excess of 1,000 U/day, anticonvulsants or phosphate-binding antacids. Finally, at least three vertebrae from L1 to L4 had to be evaluable by DXA to determine BMD in this region. Mean age: 63 years Ethnicity: Not stated (patients were from 153 centres in 34 countries in Europe, Latin America, Australia, Canada, South Africa and China) Prevalent vertebral fracture: Not stated
FRAME trial[22]	Inclusion criteria: Ambulatory postmenopausal women, 55 to 90 years of age, with a T score of −2.5 to −3.5 at the total hip or femoral. Exclusion criteria: Women who had a history of hip fracture, any severe or more than two moderate vertebral fractures, a history of metabolic bone disease or conditions affecting bone metabolism, osteonecrosis of the jaw, a 25-hydroxyvitamin D level of less than 20 ng

Trial/study	Patient characteristics
	per millilitre, current hypercalcemia or hypocalcaemia, or recent use of drugs affecting bone metabolism (within defined washout periods). Mean age: 71 years Ethnicity: Hispanic (40%); non-Hispanic (60%) Prevalent vertebral fracture: 18%
FREEDOM trial[15]	Inclusion criteria: Women between the ages of 60 and 90 years with a BMD T score of less than −2.5 at the lumbar spine or total hip were eligible for inclusion. Exclusion criteria: Women were excluded if they had conditions that influence bone metabolism or had taken oral bisphosphonates for more than 3 years. If they had taken bisphosphonates for less than 3 years, they were eligible after 12 months without treatment. Women were also excluded if they had used intravenous bisphosphonates, fluoride, or strontium for osteoporosis within the past 5 years; or parathyroid hormone or its derivatives, corticosteroids, systemic hormone-replacement therapy, selective oestrogen-receptor modulators, or tibolone, calcitonin, or calcitriol within 6 weeks before study enrolment. Mean age: 72 years Ethnicity: Not stated (the trial was conducted in Western Europe, Eastern Europe, Latin America, North America, Australia, and New Zealand) Prevalent vertebral fracture: 24%
Hadji et al. 2012[29]	Inclusion criteria: Women ≥45 years of age and at least 2 years postmenopausal were eligible if they had a history of back pain for ≥2 months before screening that was likely, in the opinion of the investigator, to be caused by osteoporotic vertebral fracture, despite conservative analgesic treatment; a baseline mean pain score of at least 4.0 on the numeric rating scale during the week before randomisation; lumbar spine, femoral neck, or total hip BMD T-score of ≤−2; and a minimum of one moderate vertebral fracture. Exclusion criteria: Exclusion criteria included diseases affecting bone metabolism other than osteoporosis; elevated serum calcium values, abnormal serum thyroid-stimulating hormone, parathyroid hormone, or 25-hydroxyvitamin D levels; imminent need for kyphoplasty or vertebroplasty; and evidence of significant pathology related to back pain which would make the interpretation of the back pain related to an osteoporotic vertebral fracture difficult, based on investigator assessment. Mean age: 71 years Ethnicity: Caucasian (80%); East Asian (0.4%); Hispanic (18%); Native American (0.4%), African Descent (0.8%) Prevalent vertebral fracture: 90%
HORIZON-PFT trial[13]	Inclusion criteria: Postmenopausal women between the ages of 65 and 89 years were eligible for inclusion if they had a BMD T score of −2.5 or less at the femoral neck, with or without evidence of existing vertebral fracture, or a T score of −1.5 or less, with radiologic evidence of at least two mild vertebral fractures or one moderate vertebral fracture.

Trial/study	Patient characteristics
	Exclusion criteria: Ineligibility criteria included any previous use of parathyroid hormone or sodium fluoride, use of anabolic steroids or growth hormone within 6 months before trial entry or oral or intravenous systemic corticosteroids within 12 months, and any previous use of strontium. Patients with a serum calcium level of more than 2.75 mmol per litre or less than 2.00 mmol per litre were ineligible, as were patients with a calculated creatinine clearance of less than 30.0 ml per minute at either of two baseline visits or urine dipstick results of more than 2+ for protein, without evidence of contamination or bacteriuria. Mean age: 73 years Ethnicity: Not stated (the trial was conducted in Western Europe, Eastern Europe, Latin America, North America, Oceania, and Asia) Prevalent vertebral fracture: 63%
Liberman et al. 1995[30]	Inclusion criteria: Women who were 45 to 80 years old and postmenopausal (≥5 years since menopause) with osteoporosis (defined as a BMD of the lumbar spine that was at least 2.5 SD below the mean value in premenopausal white women) were eligible for participation. Exclusion criteria: We excluded women with other causes of osteoporosis (e.g., treatment with glucocorticoids) or other disorders of bone and mineral metabolism (e.g., vitamin D deficiency, Paget's disease, or hyperparathyroidism); active peptic ulcer disease, abnormal renal function (serum creatinine level, > 1.5 mg per decilitre (130 mmol per litre)), or abnormal hepatic function; abnormalities of the lumbar spine precluding the assessment of BMD at a minimum of three lumbar vertebrae or a history of hip fracture; or any prior treatment with bisphosphonates or treatment within the preceding 12 months with oestrogen, progestin, calcitonin, fluoride, or an anabolic steroid. Mean age: 64 years Ethnicity: Not stated (the trial was conducted in the United States, Australia, Canada, Europe, Israel, Mexico, New Zealand, and South America) Prevalent vertebral fracture: 21%
Liu et al. 2004[31]	Inclusion criteria: Postmenopausal women between 50 and 80 years, who were free of severe or chronically disabling conditions, had their last menstrual period at least 2 years before the beginning of the study, and had a T-score for femoral neck or lumbar spine BMD measurements ≤2.5. Exclusion criteria: Known, suspected or history of carcinoma of the breast or oestrogen-dependent neoplasia, history of cancer within the previous 5 years, history of deep vein thrombosis, requirement of high-dose heparinization (>7500 U/d), bone disorders except for osteoporosis, treatment with any drug affecting bone metabolism, acute or chronic liver disease (bilirubin >34 umol/l, alanine transaminase >100 U/l, or alkaline phosphatase >300 U/l), impaired kidney function (serum creatinine >177 umol/l), or abnormal uterine bleeding of an unknown origin. Mean age: 65 years Ethnicity: Not stated (study conducted in three hospitals in China) Prevalent vertebral fracture: <=18%

Trial/study	Patient characteristics
Lufkin et al. 1998[32]	Inclusion criteria: Women with postmenopausal osteoporosis. Subjects were eligible if they were in good health except for osteoporosis, free of any serious acute or chronic medical condition that might affect bone or calcium metabolism, fully ambulatory, between the ages of 45 and 75 years, and postmenopausal (no menses for 5 years or levels of serum oestradiol <73 pmol/l and serum follicle-stimulating hormone (FSH) >30 IU/l). Exclusion criteria: Specific exclusion criteria included patients with a history of deep venous thrombosis, thromboembolic disorders, or cerebral vascular accident, also patients with a history of cancer within the previous 5 years, except for superficial skin cancer. Mean age: 68 years Ethnicity: Not stated (study conducted in two hospitals in the USA) Prevalent vertebral fracture: Not stated
MORE trial[33]	Inclusion criteria: Women who were at least 2 years postmenopausal, and who had osteoporosis, defined by BMD T-score of -2.5 or less and/or the presence of radiographically apparent vertebral fracture. Exclusion criteria: Not stated. Mean age: 74 years Ethnicity: Not stated Prevalent vertebral fracture: 37%
Morii et al. 2003[14]	Inclusion criteria: Women who were two or more years postmenopausal and no older than 80 years. All participants were Japanese who had osteoporosis, defined as L2-L4 BMD T-score of at least 2.5 SDs below the young adult mean and had a diagnosis consistent with the criteria for the diagnosis of osteoporosis in Japan. Exclusion criteria: Women were excluded from participation in the study if they had experienced bone disease other than primary osteoporosis, severe postmenopausal symptoms requiring oestrogen replacement therapy, history of or suspected breast carcinoma, any history of other cancer within the previous 5 years, except for excised superficial lesions; abnormal uterine bleeding, a history of deep venous thrombosis or thromboembolic disorders, as determined by evaluation of the participant questionnaire; endocrinologic disorders requiring pharmacologic therapy, acute or chronic hepatic disorder, with impaired kidney function (serum creatinine >225 lmol/l or >2.5 mg/dl); recent history of kidney stones; untreated malabsorption syndromes; or consumed an excess of alcohol or abused drugs. Participants were also excluded if, in the opinion of the investigator, they had pathologic fractures or if satisfactory evaluation of DXA could not be obtained due to X-ray findings. Patients were excluded if they had taken androgen, calcitonin, or bisphosphonate within the previous 6 months; been taking systemic oestrogen and progestin for up to one cycle (28 days) within the previous 6 months, or any systemic use within the previous 2 months; been taking the active form of vitamin D3, vitamin K2, or ipriflavone within the previous 3 months; been receiving fluoride therapy for more than 3 months during the previous 2 years; undergone systemic corticosteroid therapy for more than 1 month within the past year; or taken antiseizure drugs or pharmacologic doses of vitamin D. Participants who participated

Trial/study	Patient characteristics
	in other clinical trials within 4 months before registration or who had participated in any other clinical trial of raloxifene hydrochloride were also excluded. Mean age: 65 years Ethnicity: Japanese (100%) Prevalent vertebral fracture: 26%
Neer et al. 2001[34]	Inclusion criteria: Women were eligible for enrolment if they were ambulatory, if a period of at least five years had elapsed since menopause, and if they had at least one moderate or two mild atraumatic vertebral fractures on radiographs of the thoracic and lumbar spine, and an ambulatory status. For women with fewer than two moderate fractures, an additional criterion for enrolment was a value for BMD of the hip or lumbar spine that was at least one SD below the mean value in normal premenopausal white women (age range, 20 to 35 years). Exclusion criteria: Women with illnesses that affect bone or calcium metabolism, urolithiasis within the preceding 5 years, impaired hepatic function, a serum creatinine concentration exceeding 2 mg per decilitre (177 µmol per litre), or alcohol or drug abuse, as well as women who had taken drugs that alter bone metabolism within the previous 2 to 24 months (depending on the drug) were excluded. Mean age: 70 years Ethnicity: White (99%) Prevalent vertebral fracture: 100%
ROSE trial[35]	Inclusion criteria: Women aged 55 to 90 years who were considered postmenopausal based on either spontaneous amenorrhea or following surgical bilateral oophorectomy or after hysterectomy with serum FSH >20 IU/l and oestradiol <10 pg/ml. Eligible patients also had an increased risk of fracture, based on DXA T-score ≤−2.0 at total hip or spine (L1–L4) within 3 months prior to screening and clinical risk factors. Exclusion criteria: Patients who had received prior therapy with bisphosphonates, parathyroid hormone, strontium ranelate, raloxifene, calcitonin, high-dose corticosteroids, or hormone replacement within 6 months prior to randomisation; patients with a fracture within 6 months prior to randomisation, secondary osteoporosis, primary hyperparathyroidism, and presence of contraindications to study drugs were excluded. Other exclusion criteria included calculated creatinine clearance <35 mL/min; serum calcium >2.75 mmol/L, or <2.00 mmol/L; serum alkaline phosphatase higher than 2.5 times the upper limit of normal; any kind of jawbone disease or infection that may necessitate oral surgery during the course of the study and any tooth extractions during the last 3 months; or surgery of the jaw during the last 6 months before inclusion in the study. Patients with a history of invasive malignancy of any organ system within the past 5 years (excluding basal cell or squamous cell carcinoma of the skin) were also excluded. Mean age: 68 years Ethnicity: Caucasian (99%) Prevalent vertebral fracture: Not stated

Trial/study	Patient characteristics
RUTH trial[36]	Inclusion criteria: Postmenopausal women ≥55 year of age, ≥1 year postmenopausal, and had established CHD or were at high risk for CHD. Exclusion criteria: Not stated. Mean age: 68 years Ethnicity: White (84%) Prevalent vertebral fracture: Not stated
Silverman et al. 2008 (93)[37]	Inclusion criteria: Generally healthy women between the ages of 55 and 85 years were eligible for study inclusion if they were at least 2 years postmenopausal and had osteoporosis, defined as low BMD or radiographically confirmed vertebral fractures. Subjects without prevalent vertebral fracture were required to have lumbar spine or femoral neck BMD T-scores between −2.5 and −4.0 (inclusive), whereas subjects with prevalent vertebral fracture (at least one mild vertebral fracture) were required to have lumbar spine and femoral neck BMD T-scores not worse than −4.0. Exclusion criteria: Women were excluded if they had diseases that may affect bone metabolism, conditions that could interfere with bone mineral densitometry, pathologic vertebral fractures, vasomotor symptoms requiring treatment, or serious conditions such as endometrial hyperplasia or carcinoma, abnormal vaginal bleeding, malignancy within 10 years of the study, endocrine disorders requiring treatment, or untreated malabsorption disorders. Subjects with an active or history of deep vein thrombosis, pulmonary embolism, or retinal vein thrombosis were also excluded, as were subjects with elevated fasting total cholesterol or triglyceride levels (≥310 or ≥300 mg/dl, respectively). The use of androgens, systemic oestrogen (except estriol 2.0 mg/d), topical oestrogen (>3 times per week), progestogens, SERMs, bisphosphonates, calcitonin, parathyroid hormone, and cholecalciferol (>50,000 IU per week) was prohibited within 6 months of screening. Mean age: 66 years Ethnicity: White (87%) Prevalent vertebral fracture: 56%
VERO trial[38]	Inclusion criteria: Ambulatory post-menopausal women older than 45 years of age with a BMD T score less than or equal to –1.50 SDs at the femoral neck, total hip, or lumbar spine. Participants had to have radiographic evidence of at least two moderate (i.e., a reduction in vertebral body height of 26% to 40%) or one severe (more than 40% reduction) prevalent vertebral fragility fracture according to the classification of Genant and colleagues. Exclusion criteria: Patients with unresolved skeletal diseases other than osteoporosis, malignant tumours in the 5 years before screening, osteonecrosis of the jaw, previous atypical subtrochanteric femoral fractures, risk factors for osteosarcoma, GI disorders contraindicating risedronate, significantly impaired hepatic function, or a calculated creatinine clearance less than 30 mL/min using the Cockcroft–Gault equation. We also excluded patients who had undergone kyphoplasty or vertebroplasty at three or more levels before randomisation or

Trial/study	Patient characteristics
	within the 6 months before randomisation. Participants had to have normal baseline serum albumin-corrected calcium, parathyroid hormone, and free thyroxine concentrations, and 25-hydroxy-vitamin D concentration greater than 23 nmol/L. Mean age: 72 years Ethnicity: White (98%) Prevalent vertebral fracture: 100%
VERT MN trial (EU analysis)[39]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures. Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, and use of calcitonin, calcitriol or vitamin D supplements within 1-month, anabolic steroids, oestrogen, oestrogen-related drugs or progestogen within 3 months, or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months. Mean age: 71 years Ethnicity: Not stated (patients in the European analysis were all from Europe) Prevalent vertebral fracture: >50%
VERT-MN trial (AUS+EU analysis)[39]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures. Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, and use of calcitonin, calcitriol or vitamin D supplements within 1-month, anabolic steroids, oestrogen, oestrogen-related drugs or progestogen within 3 months, or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months. Mean age: 71 years Ethnicity: Not stated (patients were recruited in 80 European and Australian centres) Prevalent vertebral fracture: >50%
VERT-MN trial (NAm analysis)[40]	Inclusion criteria: Ambulatory women up to 85 years old and at least 5 years postmenopausal were eligible if they had at least two radiographically confirmed vertebral (T4–L4) fractures or one vertebral fracture and low lumbar-spine (L1-L4) BMD (defined as <-0.83 g/cm^2 (Hologic instrument) or ≤0.94 g/cm^2 (Lunar instrument)). Exclusion criteria: Exclusion criteria included conditions that might interfere with evaluation of spinal osteoporosis, or received drugs known to affect bone metabolism (e.g. calcitonin, calcitriol or cholecalciferol supplements within 1 month; anabolic steroids, oestrogen, oestrogen-related drugs or progestins within 3 months; or bisphosphonates, fluoride or subcutaneous oestrogen implant within 6 months). Mean age: 69 years Ethnicity: Not stated (patients were recruited in 110 North American centres) Prevalent vertebral fracture: 80%

Trial/study	Patient characteristics
ZONE trial[41]	Inclusion criteria: Subjects were male and female Japanese patients aged between 65 and 89 years, and were ambulatory patients who had been diagnosed with primary osteoporosis based on the Diagnostic Criteria for Primary Osteoporosis of the Japanese Society for Bone and Mineral Research; patients who have fragility fractures caused by low BMD (young adult mean <80 %; T score <−1.7), with between one and four vertebral fractures from the fourth thoracic to the fourth lumbar vertebra (Th4 to L4). Exclusion criteria: Key exclusion criteria were a history of bisphosphonate use within 2 years prior to the study; serious complications including the heart, liver, or kidney disease; creatinine clearance <35.0 mL/min or urinary protein ≥2+; serum calcium <8.0 mg/dL or >11.0 mg/dL; and undergoing or planning to undergo an invasive dental procedure of the jawbone, such as tooth extraction, at the time informed consent was obtained. Mean age: 74 years Ethnicity: Japanese (100%) Prevalent vertebral fracture: 100%

AUS = Australia; BMD = bone mineral density; CHD = coronary heart disease; DXA = dual-energy x-ray absorptiometry; EU = European Union; FSH = follicle-stimulating hormone; GI = gastrointestinal; NAm = North America; NMA = network meta-analysis; SD = standard deviation; SERM = selective oestrogen receptor modulator; SI = Système international (d'unités), English: International System of Units; USA = United States of America

Additionally, the rates of fractures were presented in the CS at different time points for all comparator treatments, including placebo. As such, it is possible to compare the fracture rates across studies for placebo, which should be similar if the populations are similar. Across all fracture types and time points, the variability in fracture rates between studies included in the same NMA were large: for new vertebral fractures, the fracture rates varied between 1.4% and 40.0% at 12 months, 3.7% and 24.6% at 24 months and 4.1% and 25.7% at 36 months; for non-vertebral fractures, the fracture rates varied between 3.0% and 11.3% at 12 months, 3.0% and 11.3% at 24 months and 4.2% and 14.4% at 36 months; and for hip fractures, the fracture rates varied between 0.2% and 1.2% at 12 months, 0.2% and 8.7% at 24 months and 0.7% and 3.9% at 36 months. While the variation in fracture rates was largest for smaller studies, larger studies also had large variation: this is problematic as we would expect smaller studies to have more variable fracture rates than larger studies, which should have much closer fracture rates if the populations were similar. This is not necessarily indicative of potential effect modification, as, so long as fracture rates in a population in the absence of treatment are not effect modifiers, differences in the fracture rates do not by themselves indicate potential bias. However, very different fracture rates for placebo arms indicate large differences between populations, and some of these differences may be between effect modifiers, leading to potentially very large and undetectable biases.

3.3.4 Risk of bias assessment of the trials in the NMAs

The RoB assessments from the company for ARCH, FRAME and STRUCTURE are presented in Table 3.20, and for all other studies in the NMAs in Table 121 of Appendix D of the CS.[8] The ERG has checked the RoB assessments for the ARCH, FRAME and STRUCTURE trials and has no concerns about these assessments. The ERG did not assess the RoB for trials providing non-romosozumab evidence in the NMAs.

Table 3.20: Quality assessment for ARCH, FRAME and STRUCTURE

Trial number (acronym)	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
Was randomisation carried out appropriately?	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	No
Were there any unexpected imbalances in drop-outs between groups?	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No
Did the analysis include an intention-to-treat analysis?	Yes	Yes	Yes

Trial number (acronym)	NCT01631214 (ARCH)	NCT01575834 (FRAME)	NCT01796301 (STRUCTURE)
If so, was this appropriate and were appropriate methods used to account for missing data?			
Based on CS, Table 120, Appendix D.[8] Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination) CRD = Centre for Reviews and Dissemination; CS = company submission			

3.4 Critique of the indirect comparison and/or multiple treatment comparison

In total, 12 NMAs were presented by the company, covering three fracture outcomes at three timepoints as well as three BMD outcomes which were not timepoint specific. As the BMD outcomes were not included in the CE model, we will restrict the critique of the indirect comparisons to the nine NMAs of fracture outcomes. We will also limit the critique to NMAs using the ITT populations, rather than the EU label populations. Furthermore, we will critique each of the NMAs separately, with reference to the population characteristics detailed in Table 3.19 above, which details the inclusion/exclusion criteria, mean age, ethnicity and prevalent vertebral fracture rate in each study. It is unclear whether age, ethnicity and prevalent vertebral fractures are effect modifiers, but in the view of the ERG, they are all plausible effect modifiers, and thus imbalances in these variables between trials may bias any analyses.

In general, apart from the potential biases from differences in effect modifiers, the ERG believes the NMAs to be well conducted.

3.4.1 New vertebral fractures

3.4.1.1 12 months

Figure 3.4 shows the network of evidence for the analysis of new vertebral fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo ([redacted])), although it is unclear if this is due to a lack of precision. The ERG asked the company to give both the direct and indirect results for all comparisons to judge whether the inconsistency estimates were imprecise or null and precise, but the company did not provide this information. However, it is still likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.4: Network of evidence for the analysis of new vertebral fractures at 12 months

Based on Figure 7 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the comparison. The FRAME study contributes the majority of evidence for any comparisons between romosozumab and comparator treatments except alendronate, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were different in all three trials. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: there was evidence of [redacted] heterogeneity for the comparison between raloxifene and placebo ([redacted]). The mean ages varied in all trials (65 to 74 years), the ethnicities varied (three trials were not international, conducted entirely within China, Japan or the USA, compared with ARCH and FRAME which were international), and the vertebral fracture rates were similar to FRAME but not ARCH. Therefore, there is a **very high risk of bias** from effect modification in the comparison between romosozumab and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Risedronate: the VERT-MN trials were conducted in Europe, Australia and North America, but had relatively similar characteristics to the ARCH trial, though higher prevalent vertebral fracture rates than FRAME. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab and risedronate.
- Teriparatide: the VERO trial compared teriparatide and risedronate, and therefore any RoB in the comparison between romosozumab and risedronate remains in this comparison, along with any RoB for the comparison between teriparatide and risedronate. The VERO trial included only

patients with prevalent vertebral fractures, and 98% of the patients were white, which is reasonably similar to the VERT-MN trials. There is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab and teriparatide.

3.4.1.2 24 months

Figure 3.5 shows the network of evidence for the analysis of new vertebral fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.5: Network of evidence for the analysis of new vertebral fractures at 24 months

Based on Figure 10 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of a RoB from effect modification between the ARCH and FIT I trials, although the FIT I trial did not report the ethnicity of patients, and there may be effect modification from unmeasured variables. The company stated that there was no evidence of inconsistency for the two closed-loops in the network (risedronate – placebo – teriparatide (P=), and teriparatide – placebo – abaloparatide (P=)).

- Zoledronate: There was evidence for heterogeneity for the comparison between zoledronate and placebo (I^2 =). However, the patient ethnicities were different between these trials (one study was conducted solely in China, one in Japan, and one was international), the mean ages of patients was markedly different (between 57 years and 74 years), though the rate of prevalent vertebral fractures was high in all studies, as in ARCH. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.

- Raloxifene: The MORE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not report the ethnicity or location of patients. Therefore, there is **a high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: the VERT-MN AUS trial was conducted in Australia, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.
- Teriparatide: there was evidence of [redacted] heterogeneity for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.1.3 36 months

Figure 3.6 shows the network of evidence for the analysis of new vertebral fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.6: Network of evidence for the analysis of new vertebral fractures at 36 months

Based on Figure 13 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of effect modification between the ARCH and FIT I trial, though the mean age of patients and rate of prevalent vertebral fractures were both lower in the Liberman 1995 trial, and the FIT I trial did not report the ethnicity of patients. There was [redacted] observed heterogeneity between the comparison of alendronate and placebo (I^2 = [redacted]).

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: There was [redacted] observed heterogeneity between the comparison of raloxifene and placebo (I^2 = [redacted]). The MORE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not report the ethnicity or location of patients. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Australia and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.2 Non-vertebral fractures

3.4.2.1 12 months

Figure 3.7 shows the network of evidence for the analysis of non-vertebral fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo ([redacted])), although it is unclear if this is due to a lack of precision and is close to statistical significance. It is likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.7: Network of evidence for the analysis of non-vertebral fractures at 12 months

Based on Figure 16 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. The FRAME study will contribute the majority of evidence for any comparisons between romosozumab and comparator treatments, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: there was evidence of [redacted] heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all 3 trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Risedronate: the VERT-MN EU trial was conducted in Europe, but had relatively similar characteristics to the ARCH trial, though higher prevalent vertebral fracture rates than FRAME. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab and risedronate.

- Teriparatide: The Neer 2001 trial did not include any Hispanic patients and only included patients with prevalent vertebral fracture. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab and teriparatide.

3.4.2.2 24 months

Figure 3.8 shows the network of evidence for the analysis of non-vertebral fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB, as the ARCH trial does not have a high RoB.

Figure 3.8: Network of evidence for the analysis of non-vertebral fractures at 24 months

Based on Figure 19 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is some evidence of a risk of bias from effect modification between the ARCH and FOSIT trials, as the FOSIT trial included younger patients than the ARCH trial (mean of 63 years vs. 74 years), and the FOSIT trial did not report the rate of prevalent vertebral fractures. The company stated that there was [redacted] evidence of inconsistency for the closed-loop in the network (risedronate – placebo – teriparatide (P=[redacted])).

- Zoledronate: The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the ARCH trial, and therefore there is **little evidence of a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.

- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: the VERT-MN EU trial was conducted in Europe, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.
- Teriparatide: there was [redacted] observed heterogeneity for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.2.3 36 months

Figure 3.9 shows the network of evidence for the analysis of non-vertebral fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. As there was no data for non-vertebral fractures at 36 months in the ARCH trial, data from 30 months was used instead, which may have caused bias. It is possible, therefore, that the comparison between romosozumab/alendronate and alendronate has some bias, which will propagate to all other comparisons.

Figure 3.9: Network of evidence for the analysis of non-vertebral fractures at 36 months

Based on Figure 22 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis.

There is, however, little evidence of a RoB from effect modification between the ARCH and FIT I+II trials, though the FIT trials did not report the ethnicity of patients.

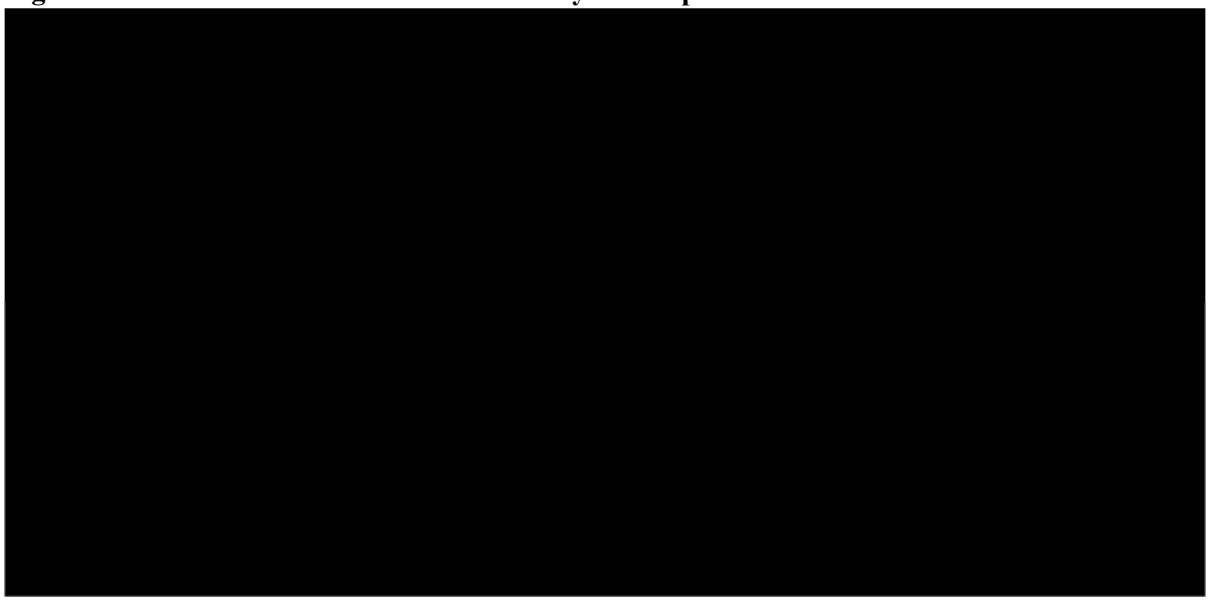
- Zoledronate: There was [redacted] observed heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: there was [redacted] observed heterogeneity between the comparison of raloxifene and placebo (I^2 = [redacted]). The RUTH was relatively similar to the ARCH trial. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: there was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Europe and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.3 Hip fractures

3.4.3.1 12 months

Figure 3.10 shows the network of evidence for the analysis of hip fractures at 12 months. This network shows that the ARCH and FRAME trials provide direct evidence for the comparisons of romosozumab and alendronate, and romosozumab and placebo. The company states that there was no evidence of inconsistency in the closed loop in the network (romosozumab – alendronate – placebo [redacted]), although it is unclear if this is due to a lack of precision. It is therefore likely that comparisons between romosozumab, alendronate and placebo do not have high RoB.

Figure 3.10: Network of evidence for the analysis of hip fractures at 12 months



Based on Figure 25 of Appendix D of the CS.[8]
CS = company submission

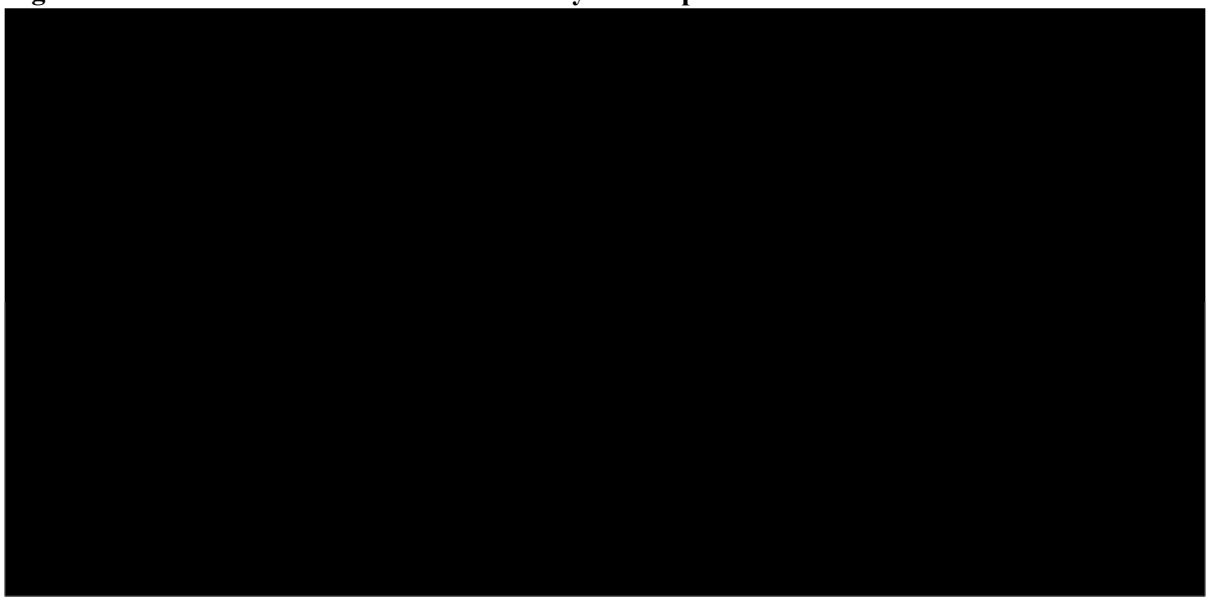
All other comparisons with romosozumab are indirect, passing through placebo. Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. The FRAME study will contribute the majority of evidence for any comparisons between romosozumab and comparator treatments, as the precision of indirect estimates decreases as more treatments are passed through: the ARCH trial must pass through alendronate before getting to placebo, so will contribute substantially less information that the FRAME trial.

- Zoledronate: there was [redacted] observed heterogeneity for the comparison of zoledronate and placebo (I^2 = [redacted]). Horizon-PFT trial had a similar mean age to ARCH and FRAME, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all three trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the FRAME trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the FRAME trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.

3.4.3.2 24 months

Figure 3.11 shows the network of evidence for the analysis of hip fractures at 24 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. It is likely that comparison between romosozumab/alendronate and alendronate does not have a high RoB as the ARCH trial does not have a high RoB.

Figure 3.11: Network of evidence for the analysis of hip fractures at 24 months



Based on Figure 28 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence of effect modification between the ARCH and FIT I trials, although the FIT I trial did not report the ethnicity of patients, and there may be effect modification from unmeasured variables. The company stated that there was [redacted] evidence of inconsistency for the closed-loop in the network (teriparatide – placebo – abaloparatide [P=[redacted]]).

- Zoledronate: There was [redacted] evidence of heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH, but patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH was 32% Hispanic and 68% non-Hispanic, and the prevalent vertebral fracture rate was lower in Horizon-PFT. The Bai 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.
- Raloxifene: the RUTH trial was relatively similar to the ARCH trial, and therefore there is **little evidence for a risk of bias** from effect modification, though unobserved and unmeasured effect modifiers may still bias this comparison.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was evidence of [redacted] heterogeneity between the comparison of teriparatide and risedronate (I^2 = [redacted]). The VERO trial compared teriparatide and risedronate, and therefore any RoB in the comparison between romosozumab and teriparatide remains in this comparison, along with any RoB for the comparison between teriparatide and risedronate. In the VERO trial, 98% of the patients were white, while in ARCH 32% of patients were Hispanic and 68% of patients were non-Hispanic. The Hadji 2012 trial had similar patient

characteristics as the ARCH trial. There is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.

- Teriparatide: the company state that [redacted] heterogeneity was observed for the comparison between teriparatide and placebo (I^2 = [redacted]). The Neer 2001 and ACTIVE trials were relatively similar to the ARCH trial, though the ACTIVE trial had a lower rate or prevalent vertebral fractures and neither trial included any Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and teriparatide.
- Abaloparatide: The ACTIVE trial had a lower rate of prevalent vertebral fractures than the ARCH trial and did not include Hispanic patients. There is therefore a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and abaloparatide.

3.4.3.3 36 months

Figure 3.12 shows the network of evidence for the analysis of hip fractures at 36 months. This network shows that the ARCH trial alone provides direct evidence for the comparisons of romosozumab/alendronate and alendronate. As there was no data for hip fractures at 36 months in the ARCH trial, data from 30 months was used instead, which may have caused bias. It is possible, therefore, that the comparison between romosozumab/alendronate and alendronate has some bias, which will propagate to all other comparisons.

Figure 3.12: Network of evidence for the analysis of hip fractures at 36 months

Based on Figure 31 of Appendix D of the CS.[8]
CS = company submission

All other comparisons with romosozumab/alendronate are indirect, passing through both alendronate and placebo, markedly increasing the uncertainty of these comparisons compared with the same comparisons at 12 months (since the FRAME study only provides evidence up to 12 months). Any imbalances in effect modifiers between studies used in the indirect comparison will bias the analysis. There is little evidence for a RoB from effect modification between the ARCH and FIT I trials, and the FIT I trial did not report the ethnicity of patients.

- Zoledronate: There was [redacted] observed heterogeneity between the comparison of zoledronate and placebo (I^2 = [redacted]). The Horizon-PFT trial had a similar mean age to ARCH and FRAME, but

patients are unlikely to be of a similar ethnic background as Horizon-PFT was conducted in Europe, Latin and North America, Oceania and Asia, while the ethnicity of patients in ARCH and FRAME were 32% and 40% Hispanic and 68% and 60% non-Hispanic respectively, and the prevalent vertebral fracture rates were very different in all 3 trials. The Chao 2013 trial included much younger patients solely from China. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and zoledronate.

- Raloxifene: The RUTH was relatively similar to the ARCH trial. The Silverman 2008 trial had younger patients with a lower rate of prevalent vertebral fractures and greater percentage of patients had white ethnicity. Therefore, there is a **high risk of bias** from effect modification in the comparison between romosozumab/alendronate and raloxifene.
- Denosumab: the FREEDOM trial was relatively similar to the ARCH trial, though the rate of prevalent vertebral fractures was lower. Therefore, there is likely a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and denosumab.
- Risedronate: There was [redacted] observed heterogeneity between the comparison of risedronate and placebo (I^2 = [redacted]). The VERT-MN trials were conducted in Europe and North America, rather than internationally for ARCH, but had relatively similar other characteristics. As such, there is a **moderate risk of bias** from effect modification in the comparison between romosozumab/alendronate and risedronate.

3.4.4 Summary

Overall, there was little direct evidence for comparisons for romosozumab included in any of the NMAs, and most studies had differences in mean age, ethnicity or rate of prevalent vertebral fractures, indicating at least a moderate RoB from effect modification. Additionally, as almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone.

Additionally, individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the RoB. As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low RoB; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high RoB.

3.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work on clinical effectiveness was undertaken by the ERG.

3.6 Conclusions of the clinical effectiveness section

The decision problem is largely in line with the NICE scope. However, the population in the CS is postmenopausal women with severe osteoporosis at high risk of fracture; where "high risk of fracture" is defined as having suffered a fracture within the last 2 years (also revert to as "imminent risk of fracture").[1] This is narrower than the population in the NICE scope (Postmenopausal women with severe osteoporosis at high risk of fracture; where timing of previous fracture is not mentioned),[2] and

narrower than the population in the ARCH trial (Postmenopausal women with severe osteoporosis at high risk of fracture; where timing of fracture is not an inclusion criterion for some patients).[3]

The clinical effectiveness evidence for romosozumab in severe osteoporosis in the CS is mainly based on the ARCH trial. Two other phase III clinical trials, the FRAME and STRUCTURE trials are mentioned in the CS as well. However, neither the FRAME nor STRUCTURE trials studied a patient population aligned to where the company expects romosozumab to be used in National Health Service (NHS) clinical practice.[17, 18] A fourth study, the BRIDGE study, considered use in men, which is not part of the marketing authorisation for romosozumab.[19]

The ARCH trial is a phase III, multicentre, randomised, double-blind trial, comparing romosozumab followed by alendronate vs. alendronate alone in postmenopausal women with severe osteoporosis and a fragility fracture (see Table 3.5).[3] This trial provides evidence for romosozumab in its expected position in the clinical pathway: a first-line therapy in patients who have previously suffered a MOF. Efficacy outcomes reported in ARCH include incidence of clinical, vertebral, non-vertebral and hip fracture and percentage change from baseline in BMD. Data from ARCH were used as the main data for the economic modelling in this submission.

In the ARCH trial, romosozumab/alendronate statistically significantly reduced the incidence of new vertebral fractures at month 24, meeting its primary endpoint. Patients in the romosozumab/alendronate arm had a 50% lower relative risk of vertebral fractures compared to patients on alendronate alone over 24 months (RR 0.50, 95% CI 0.38 to 0.66).[20] Additionally, a statistically significantly lower proportion of patients experienced a clinical fracture (non-vertebral fracture and clinical vertebral fracture) at the time of primary analysis in the romosozumab/alendronate group compared to alendronate alone (HR 0.73, 95% CI 0.61 to 0.88), meeting the other primary endpoint.[20] At the primary analysis there were also a lower number of patients who experienced non-vertebral fractures (HR 0.81, 95% CI 0.66 to 0.99) and hip fractures (HR 0.62, 95% CI 0.42 to 0.92).[20] Patients treated with romosozumab also had a statistically significantly greater increase in BMD from baseline compared to alendronate (adjusted $P<0.001$), which was maintained until month 36.[20]

The Kaplan-Meier curves for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3) show that there is a visible separation of the romosozumab/alendronate and alendronate arms in terms of time to first fracture up to month 42. However, the curves seem to converge again by month 48. This means that it is possible that the effects of romosozumab wane over time. However, by 48 months the curves are based on smaller numbers of patients which increases uncertainty. Therefore, longer term follow-up is needed to see whether the effects are maintained over time.

Overall, results of the ARCH trial are favourable for romosozumab. Both primary outcomes (the cumulative incidence of new vertebral fracture at month 24 and the cumulative incidence of clinical fracture at time of primary analysis) are met and most fracture results significantly favour romosozumab over alendronate. In addition, all BMD outcomes significantly favour romosozumab over alendronate. However, the graphs for time to first clinical fracture (Figure 3.2) and time to first non-vertebral fracture (Figure 3.3), seem to indicate that the effectiveness of romosozumab over alendronate becomes less after 42 months; longer term follow-up is needed to see whether the effects are maintained over time.

The incidences of AEs and SAEs were similar overall in the ARCH trial between the two treatment groups during the 12-month double-blind period, and cumulative incidences were similar between the two groups during the primary analysis period. However, more people in the romosozumab group

experienced adjudicated serious CV AEs during the double-blind period, with 50 patients (2.5%) in the romosozumab group and 38 (1.9%) in the alendronate group reporting these events (odds ratio (OR) 1.31, 95% CI 0.85 to 2.00). A total of 16 patients (0.8%) in the romosozumab group and 6 (0.3%) in the alendronate group reported cardiac ischemic events (OR 2.65; 95% CI, 1.03 to 6.77), and 16 patients (0.8%) in the romosozumab group and 7 (0.3%) in the alendronate group reported cerebrovascular events (OR 2.27, 95% CI 0.93 to 5.22). Therefore, romosozumab is contraindicated for patients with a history of MI or stroke.

The company claims that it is *"reasonable to conclude that a population of patients treated with romosozumab will experience a reduced level of pain, disability and mortality, relative to patients treated with currently available treatments, because these patients will experience fewer fragility fractures compared to patients treated with currently available treatments"* (Response to request for clarification, question A11).[9] However, after 12 months, more patients died in the romosozumab group (n=30, 1,5%) than in the alendronate group (n=21, 1.0%). At the time of the primary analysis, 90 patients had died in both groups.

In total, 11 NMAs were presented by the company, covering five distinct outcomes at three timepoints, although only the three fracture outcomes were used in the CE model. In these NMAs, many comparator treatments were directly and indirectly compared with romosozumab using Bayesian methods. The methods used appear valid and appropriate.

However, there was little direct evidence for comparisons for romosozumab included in any of the NMAs, and most studies had differences in mean age, ethnicity or rate of prevalent vertebral fractures. As these variables could potentially be effect modifiers when conducting indirect comparisons, different levels of these variables in the included studies likely indicates at least a moderate risk of bias from effect modification. Additionally, as almost all comparisons did not include direct evidence, inconsistency could only rarely be assessed, and as most direct comparisons only included a single study, heterogeneity could also only rarely be assessed. This is particularly problematic as the direct evidence for romosozumab came from only two trials (FRAME and ARCH), which did not have the same comparators, and the FRAME trial only provided data up to 12 months. Therefore, almost all evidence in this submission comes from the ARCH study alone.

Additionally, individual studies rarely provided data consistently across timepoints, and some studies that were missing data at one timepoint had data from an earlier timepoint used instead (e.g. the ARCH study did not have data at 36 months for non-vertebral fractures, so used data from 30 months instead). There were also large differences in the rates of fractures in the placebo arms of different studies, indicating large differences in the populations that likely extend to unknown and unmeasured effect modifiers, increasing the RoB. As such, only the comparisons between romosozumab, alendronate and placebo can be considered to have a low RoB; all other comparisons are indirect and most commonly have observed differences in variables likely to be effect modifiers, and therefore, when considered across all timepoints and outcomes, almost all are considered to have a high RoB.

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

One set of systematic literature searches was performed to identify CE studies and costs and healthcare resource use studies (CS Appendix G and Appendix I).[8] Searches were not conducted to identify health-state utility values. Instead, economic evaluations included in the original and update economic systematic literature reviews (SLRs) were reviewed for novel health-state utility values of relevance to the CE model for romosozumab.

4.1.1 Searches performed for cost effectiveness section

Appendices G and I of the CS reported the literature searches used to identify CE studies and costs and healthcare resource use studies.[8] Searches were conducted in March and April 2018 and an update search was conducted in February and March 2021. Summaries of the resources searched are provided in Tables 4.1 and 4.2. The following paragraphs contain summaries and critiques of all searches related to CE presented in the CS.

Table 4.1: Resources searched for cost effectiveness studies and costs and healthcare resource use studies. March/April 2018

Resource	Host/Source	Date Range	Date searched
Databases			
Embase	OvidSP	1974 to 9th March 2018	9 March 2018
MEDLINE, including MEDLINE Daily, MEDLINE In-Process and Epub Ahead of Print	OvidSP	1946 to 9th March 2018	9 March 2018
NHS EED	Cochrane Library: Wiley Online	Issue 2 of 4, April 2015	9 March 2018
HTA Database	Cochrane Library: Wiley Online	Issue 4 of 4, October 2016	9 March 2018
EconLit	EBSCO	1886 to 8th March 2018	9 March 2018
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear4/	-	13 April 2018
ScHARRHUD	www.scharrhud.org/	-	13 April 2018

Resource	Host/Source	Date Range	Date searched
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/	-	13 April 2018
Conference Proceedings			
WCO-IOF-ESCEO	PDF abstract books	2016 and 2017	13 April 2018
ECTS	PDF abstract books	2016 and 2017	13 April 2018
ASBMR	PDF abstract books	2016 and 2017	13 April 2018
ISPOR	https://www.ispor.org/heor-resources/presentations-database/search	2016 and 2017	13 April 2018
FFN	PDF abstract book	2016 and 2017	13 April 2018
EULAR	http://scientific.sparx-ip.net/archiveeular/	2016 and 2017	13 April 2018
HTA websites			
NICE	https://www.nice.org.uk/	-	13 April 2018
SMC	https://www.scottishmedicines.org.uk/	-	13 April 2018
AWMSG	www.awmsg.org/	-	13 April 2018
NCPE	http://www.ncpe.ie/	-	13 April 2018

The bibliographies of all relevant SLRs, meta-analyses and HTA submissions identified through the electronic database and HTA agency website searches were also manually searched to identify any additional studies of relevance.

ASBMR = American Society for Bone and Mineral Research; AWMSG = All Wales Medicines Strategy Group; CEA = cost effectiveness analysis; ECTS = European Calcified Tissue Society; EED = Economic Evaluation Database; EQ-5D = EuroQol-5 Dimensions; EULAR = European League Against Rheumatism; FFN = Fragility Fracture Network; HTA = health technology assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NCPE = National Centre for Pharmacoeconomics; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; ScHARRHUD = School of Health and Related Research Health Utilities Database; SLR = systematic literature review; SMC = Scottish Medicines Consortium; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases

Table 4.2: Resources searched for cost effectiveness studies and costs and healthcare resource use studies. February/March 2021

Resource	Host/Source	Date Range	Date searched
Databases			
Embase	OvidSP	1974 to 24th February 2021	24 February 2021
MEDLINE, including MEDLINE Daily, MEDLINE In-Process and Epub Ahead of Print	OvidSP	1946 to 24th February 2021	24 February 2021
NHS EED	Centre for Reviews and Dissemination	Issue 2 of 4, April 2015	24 February 2021
HTA Database	Centre for Reviews and Dissemination	Issue 4 of 4, October 2016	24 February 2021
INAHTA HTA Database	Not reported	from 1996 to 24th February 2021	24 February 2021
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear4/	-	5 March 2021
ScHARRHUD	www.scharrhud.org/	-	5 March 2021
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/	-	5 March 2021
Conference Proceedings			
WCO-IOF-ESCEO	PDF abstract books	2019 and 2020	5 March 2021
ECTS	PDF abstract books	2019 and 2020	5 March 2021
ASBMR	PDF abstract books	2019 and 2020	5 March 2021
ISPOR	https://www.ispor.org/heor-resources/presentations-database/search	2019 and 2020	5 March 2021
FFN	PDF abstract book	2019	5 March 2021
EULAR	http://scientific.sparx-ip.net/archiveeular/	2019 and 2020	5 March 2021

Resource	Host/Source	Date Range	Date searched
HTA websites			
NICE	https://www.nice.org.uk/	-	5 March 2021
SMC	https://www.scottishmedicines.org.uk/	-	5 March 2021
AWMSG	www.awmsg.org/	-	5 March 2021
NCPE	http://www.ncpe.ie/	-	5 March 2021
The bibliographies of all SLR or (network) meta-analyses ([N]MAs) identified in the course of this update were hand-searched in order to identify any additional, relevant studies for inclusion. ASBMR = American Society for Bone and Mineral Research; AWMSG = All Wales Medicines Strategy Group; CEA = cost effectiveness analysis; ECTS = European Calcified Tissue Society; EED = Economic Evaluation Database; EQ-5D = EuroQol-5 Dimensions; EULAR = European League Against Rheumatism; FFN = Fragility Fracture Network; HTA = health technology assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NCPE = National Centre for Pharmacoeconomics; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; ScHARRHUD = School of Health and Related Research Health Utilities Database; SLR = systematic literature review; SMC = Scottish Medicines Consortium; WCO-IOF-ESCEO = World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases			

ERG comment:

- The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported.
- Conference proceedings were searched. Full details of the conference searches, including search terms, URLs, results and the date of the searches, were provided. A full explanation for the two-year date limit was provided.
- Additional health economic specific resources were searched, and full details of the search strategies or search terms used, dates of searches, and results, were reported in the CS.[1, 8]
- Health technology assessment organisation websites were searched, and full details of the search terms used, dates of searches, and results, were reported in the CS.[1, 8]
- Extensive use of truncation, proximity operators, synonyms, and subject headings (MeSH and EMTREE) were included in the search strategies. Study design search filters for CE evaluations and UK cost studies were included. It would have been helpful if the search filters had been cited in the methods section.[72] There were no language or date limits.
- Update searches were conducted in February and March 2021. Full details of the searches were provided.
- Searches of NHS EED and the HTA database for the original review searches were conducted via the Cochrane Library. These resources were no longer available via the Cochrane Library by the time of the update searches in February 2021, so the CS translated the searches to run in the Centre for Reviews and Dissemination (CRD) interface. In addition, the company searched the International Network of Agencies for Health Technology Assessment (INAHTA) Database to retrieve more up-to-date health technology assessment reports. A full explanation for these changes was provided in the CS.[1, 8]
- No searches were conducted to identify health-state utility values. The CS reported that *"To supplement the search for economic data, all economic evaluations included in the original and*

update SLRs were reviewed for novel health-state utility values of relevance to the cost-effectiveness model for romosozumab. The economic evaluations were reviewed by two independent reviewers and their results compared to reach consensus. Any disagreements were resolved by a third independent reviewer, if necessary." The company did search health utilities resources (CEA Registry, ScHARRHUD and EQ-5D Publications Database).

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on CE studies, utilities and costs and resource use are presented in Table 4.3.

Table 4.3: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	Men and/or postmenopausal women with osteoporosis at increased risk of fracture. Patients may be stated to be at 'risk of fracture' in the paper, or may have been defined as at risk by the presence of at least one of the following: • Age ≥65 years (women) and ≥75 years (men) • BMD T-score of ≤2.5 • Prior fracture • Family history • Long periods of inactivity	• Patients being studied for the prevention or treatment of glucocorticoid-induced osteoporosis • Patients with normal or unspecified BMD who have not been selected based on the presence of risk factors (see left) • Patients with other indications for osteoporosis treatment, including: • Hormonal disorders, e.g., hyperthyroidism, pituitary gland disorders, Cushing's syndrome, hypogonadism • Paget's disease • Hypercalcaemia of malignancy • Breast cancer • Prostate cancer • Rheumatoid arthritis • Coeliac disease • Crohn's disease • Eating disorders, e.g., bulimia or anorexia • Heavy smoking or drinking Where studies included a mixed population of participants in which the above eligibility criteria were not met by all patients, the study was excluded unless separate data on the outcomes of interest were reported for the population of interest.
Intervention (economic evaluations)	Romosozumab, or any of the below interventions: • Teriparatide • Bisphosphonates (alendronate, risedronate, ibandronate and zoledronic acid)	• Combination therapies (with the exception of combination of an intervention of interest with vitamin D and calcium supplementation)

	Inclusion criteria	Exclusion criteria
	• Denosumab • Raloxifene • Strontium ranelateb • Abaloparatideb	• Interventions of interest that were co-administered with any other therapy, with the potential to augment bone, unless concomitant treatments were specified in the summary of product characteristics • Interventions that were not administered in accordance with their licensed indication
Intervention (cost and resource use)	Any or none	Not applicable.
Comparator	Any or no comparator	Not applicable.
Outcomes(s) 1 (Published economic evaluations)	Outcomes of relevant study designs, including: • Costs, including cost per fracture event avoided • Life years gained • Quality-adjusted life years • Number of fractures • Number of patients with fractures • Incremental costs and QALYs • Incremental cost effectiveness ratios	Studies not reporting relevant outcomes
Outcomes(s) 3 (Cost/resource use studies)	Original direct costs or resource use data published in 2008 onwards relevant to an economic model of romosozumab in the prevention of fractures in osteoporosis, including but not necessarily limited to: • Treatment and management of fractures, including: • Fractures of the hip and vertebrae • Nursing home/long-term care • BMD measurement • Physician visits • Proton pump inhibitor for gastrointestinal events • IV injections of zoledronate and denosumab • Nurse visit • Distribution of patients among treatment sites, including: • Hospital (inpatient and outpatient) • Accident and emergency department • Nursing home Data must be relevant to the UK NHS and Personal and Social Services	Studies not reporting relevant outcomes, or reporting indirect costs only

	Inclusion criteria	Exclusion criteria
Study design 1 (Economic evaluations)	Original economic evaluations considering both the costs and benefits of alternative interventions. Specifically, the following types of analysis: • Cost effectiveness • Cost utility • Cost benefit • Cost minimisation • Cost consequence To be eligible, models needed to be novel with a base-case in the UK, US, Australia or Canada. Non-novel models were only eligible if the base-case was the UK.	• Publications without original data • Study protocol reporting no results • Comments • Letters • Editorials • Non-systematic/narrative reviews • Animal/in vitro studies
Study design 3 (Cost/resource use studies)	Primary research publications on any study design	• Publications without original data • Study protocol reporting no results • Comments • Letters • Editorials • Non-systematic/narrative reviews • Animal/in vitro studies
Publication type (economic evaluations)	• Journal articles presenting original research • HTAs presenting primary research • Original SLR: Congress abstracts published in or after 2016 • During SLR update: Congress abstracts published in or after 2019	Other publications types
Publication type (cost and resource use)	• Journal articles presenting original research • SLRs of relevant primary publications (these were included at the title/abstract review stage and were used for the identification of any additional primary studies not identified through the database searches. They were excluded during the full-text review unless they reported primary, original research themselves) • HTAs presenting primary research • Original SLR: Congress abstracts published in or after 2016 • During SLR update: Congress abstracts published in or after 2019	

	Inclusion criteria	Exclusion criteria
Other (Economic evaluations)	• English language only • Human subjects only	• Articles not in the English language • Studies not in human subjects
Other (cost and resource use)	• Studies conducted in the UK • English language only • Human subjects only	• Articles not in the English language • Studies not conducted in the UK • Studies not in human subjects

Based on Tables 142 and 156 from the Appendices of the CS.[8]
[a] If a study did not specifically state that women were postmenopausal, then it was not excluded. However, if a study specifically stated that patients were not postmenopausal, it was excluded; [b] Strontium ranelate and teriparatide were included as potentially relevant comparators at the time of the original SLR, which was conducted before the NICE Scope was released.
BMD = bone mineral density; CS = company submission; HTA = health technology assessment; ICER = incremental cost effectiveness ratio; IV = intravenous; LYG = life years gained; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal and Social Services; QALY = quality-adjusted life year; SLR = systematic literature review; UK = United Kingdom; US = United States

In total, 3,732 unique articles were reviewed at the title/abstract review stage in the economic evaluation SLR.[8] Of these, 352 articles were deemed potentially relevant and reviewed at the full-text stage, with 29 articles ultimately meeting the economic evaluation inclusion criteria and three meeting the cost/resource use criteria. An additional nine articles were identified through congress searching, website searching and through handsearching of bibliographies in the economic evaluation SLR, resulting in a total of 38 articles reporting on 35 unique studies being included. These studies are summarised in Tables 143 and 147 of the CS appendices.[8] No additional cost and resource use articles were identified, resulting in a total of three studies being included in this review. These studies are summarised in Table 157 of the CS appendices.[8]

An additional SLR for HRQoL was not conducted. All economic evaluations included in the original and updated SLRs were reviewed for novel health-state utility values of relevance to the CE model for romosozumab. The handsearching of included economic evaluations did not identify any novel health-state utility values of relevance to the romosozumab model.

4.1.3 Conclusions of the cost effectiveness review

The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported. Overall, the ERG does not have any major concerns regarding the searches but notes that no searches were conducted to identify health-state utility values (see Section 4.1.1 for more details), it is unclear whether empirical studies estimating utility values in this condition were missed as only included economic evaluations were searched for utility values. Furthermore, it is unclear whether relevant resource use data were missed by including only studies conducted in the UK. Resource use data from other countries could have been considered, with UK unit costs applied.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.4: NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers.	As per the reference case.
Perspective on costs	NHS and PSS.	As per the reference case.
Type of economic evaluation	Cost utility analysis with full incremental analysis.	As per the reference case.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared.	As per the reference case.
Synthesis of evidence on health effects	Based on systematic review.	As per the reference case.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	As per the reference case.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers.	Utility multipliers for fracture events were estimated from patient reported data from the ICUROS study. These multipliers were applied to UK general population EQ-5D norms.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population.	Although not explicitly stated, it seems that the UK EQ-5D valuation tariff has been used to estimate the multipliers. The UK value set was used to the was used to estimate the general population norms.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	As per the reference case.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS.	As per the reference case.
Discounting	The same annual rate for both costs and health effects (currently 3.5%).	As per the reference case.

Element of health technology assessment	Reference case	ERG comment on company's submission
ERG = Evidence Review Group; HRQoL = health related quality of life; ICUROS = International Costs and Utilities Related to Osteoporotic Fractures Study; NHS = National Health Service; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom		

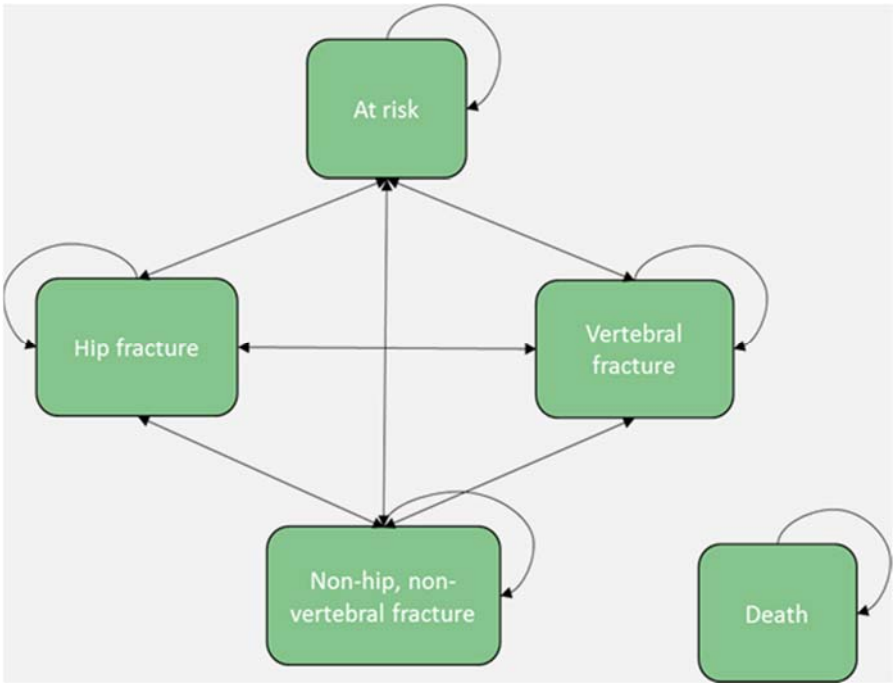
4.2.2 Model structure

4.2.2.1 Health states/events and transitions

A "*de novo*" Markov microsimulation model was developed in Microsoft Excel to assess the CE of romosozumab followed by alendronate compared to alendronate alone in postmenopausal women who have experienced a MOF within the past 24 months.

The model, shown in Figure 4.1, consisted of five health states: at risk, hip fracture, vertebral fracture, NHNV fracture and death.

Figure 4.1: Model structure



Based on Figure 13 in the CS.[1]
CS = company submission

At the start of the model, all patients are in the "at risk" health state. At the end of each model cycle patients can either transition to one of the fracture states, stay in the same health state without having a new fracture, or die. Upon transitioning to "death", patients remain there for the rest of the simulation. No restrictions were imposed for the sequence or number of fractures experienced.

As an advantage of the micro-simulation approach, the model keeps track of each patient's history to enable the calculation of costs, quality of life, and fracture risk over the lifetime (with a maximum of 100 years) of each individual patient.

At any point in the model, the risk of sustaining a fracture is based on a combination of four components:

1. The general population risk of fracture.

2. The increased fracture risk associated with osteoporosis, relative to the general population.
3. The increased fracture risk due to having sustained a recent fracture (i.e., the imminent fracture risk).
4. The reduction in risk, where applicable, due to osteoporosis treatment.

The input values, and their underlying assumptions, for each of these components are further elaborated in Section 4.2.6 of the ERG report.

The same model, but with different input values, was also used as the basis for two recent publications in the peer-reviewed journal of the International Osteoporosis Foundation: 'Osteoporosis International'.[73, 74] In Söreskog et al. 2021a the CE of romosozumab followed by alendronate compared to alendronate alone for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture was assessed from a Swedish perspective with model inputs for treatment effectiveness based on ARCH.[73] In Söreskog et al. 2021b the CE of a ("hypothetical") bone-forming agent followed by an anti-resorptive therapy compared to an anti-resorptive therapy alone was assessed for the prevention of fractures in patients with osteoporosis from a UK perspective.[74]

ERG comment: The model structure appears appropriate. However, the ERG's ability to step through and evaluate the model functionality was hindered by the fact that all model calculations are done in background VBA code. The VBA code is password protected and the company were unable to make the password available to the ERG due to confidentiality issues with the FRAX algorithm that was implemented in the VBA code. Outside of the VBA code only input parameters and hardcoded results are available. At clarification, the company did provide some of the VBA code in separate files but the ERG was unable to:

- Verify that this matched the code within the model.
- Step through the code as they would in the model to understand the functionality of the code.
- Make any changes to the code in response to potential errors or to make ERG or base-case changes (beyond changes to the available input parameters).

At a later stage a version of the model was made available to the ERG in which the VBA code was separated in a non-password protected version for the code that was not related to FRAX and a version with the password protected FRAX algorithm. However, the ERG was advised not to use this version of the model for running analyses. Therefore, the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. This means that the ERG has not been able to carry out its usual level of investigation and has had to proceed by assuming that the model functions correctly and as reported by the company.

The ERG comment in Section 5.3 presents some inconsistencies and issues found in the model and the VBA code. These appear to have a minor impact on the results, but this needs further confirmation from the company.

4.2.3 Population

The population in the Final Scope by NICE is defined as "Postmenopausal women with severe osteoporosis at high risk of fracture", in line with the marketing authorisation by the European Medicines Agency (EMA) for the use of romosozumab in women who have been through the menopause and who have severe osteoporosis (low bone density and previous fracture), leading to a high risk of further fractures. Severe osteoporosis is defined, according to the World Health Organization (WHO), based on a BMD value below a T-score of −2.5 and with one or more fragility

fractures (i.e., low impact fractures sustained from standing height or less). Importantly, the NICE final scope does not define "high risk of fracture".

The modelled population in the CS consisted of postmenopausal women with baseline characteristics, provided in Table 4.5, that are the average of those in the trial population in ARCH in terms of age (i.e., 74 years), femoral neck BMD T-score (i.e., -2.90) and BMI (i.e., 25.41). The inclusion criteria used in ARCH are listed in Section 3.2.1 of the ERG report. As described in that Section as well, the modelled population in the CS is assumed to consist of patients who have had a MOF within the prior 24 months. Based on the FRAX algorithm in combination with the additional risk that is associated with a recent fracture, the modelled population had an estimated mean 10-year MOF probability of 30%. An important difference between the ARCH ITT population and the modelled population is that ARCH included patients who previously sustained a fracture regardless of recency, whereas for the modelled population it is assumed that a previous fracture was sustained within 24 months prior to the start of treatment. In the ARCH ITT population, [redacted] of patients suffered a MOF within 24 months prior to randomisation.

Table 4.5: Baseline patient characteristics used in the economic model

Model parameter	Value	Source and appropriateness for modelling patient population in decision problem
Sex	Female	Licensed indication
Fracture history	Recent fracture (MOF within 24 months)	ARCH,[3] Swedish registry.[75] Specifying MOF aligns with the expected target population for romosozumab in clinical practice, to maximise the benefits of treatment
Mean age, years	74	ARCH[3]; comparable to the average age of postmenopausal women with osteoporosis in the UK [11, 76]
Mean femoral neck T-score (SD)	−2.90	ARCH[3]
Mean BMI	25.41	ARCH[3]
Mean 10-year MOF probability	30%	Target patient population

Based on Table 17 of the CS.[1]
BMI = body mass index; CS = company submission; MOF = major osteoporotic fracture; SD = standard deviation; UK = United Kingdom.

ERG comment: The issues with the population explained in Sections 2.1 and 3.2.1 are also applicable to the CE analyses.

4.2.4 Interventions and comparators

The modelled intervention consisted of a once-in-a-lifetime, 12-month course of romosozumab, followed by a 48-month course of alendronate. Romosozumab is administered monthly at a dose of 210 mg via two subcutaneous injections of 105 mg each into the abdomen, thigh, or upper arm. Alendronate is administered orally at a weekly dose of 70 mg.

The comparators that were used in the company base-case model consists of a 60-month course of alendronate, administered orally at a weekly dose of 70 mg, and no treatment. Additionally, the company performed a series of scenario analyses for which the following comparators were used i.e., instead of alendronate: teriparatide, denosumab, risedronate, zoledronate, and raloxifene. Teriparatide

is administered daily at a dose of 20 µg (i.e., microgram) via subcutaneous injection into the abdomen or thigh, over the course of (maximally) 24 months per lifetime. Denosumab is administered once every 6 months at a dose of 60 mg via a single subcutaneous injection into the thigh, abdomen or upper arm. Risedronate is administered orally once per week at a dose of 35 mg. Zoledronate is administered once per year via intravenous infusion at a dose of 5 mg. Raloxifene is administered orally at a daily dose of 60 mg. For all modelled comparators in the scenario analyses a treatment duration of 60 months was assumed, except teriparatide for which the maximum treatment duration of 24 months was assumed. A description of all the included treatment sequences, their durations and their residual effects is provided in Table 4.14 in Section 4.2.6.3.

ERG comment: The treatments that were used as comparators in the company's base-case and scenario analyses include all that were listed in the NICE scope, except for ibandronic acid. The company indicated that no trials for ibandronate at the licensed dose were found to be included in the NMA for fracture outcomes, and therefore this comparator was not included.

The ERG notes that there is uncertainty regarding the appropriateness and relevance of the included comparators, due to the uncertainty regarding the relevant population as described in the previous section. This is because risk of fracture is often used to guide choice of treatment.

For the information summarised above, the ERG noted some small inconsistencies in the information that was provided in Table 31 of the CS relative to information provided in the corresponding summaries of product characteristics and other general sources regarding medicines that can be found online.[1] Specifically, the ERG added the daily dose of teriparatide and corrected the dosage of zoledronate (5 mg instead of 4 mg) and frequency of administration for denosumab (once every 6 months instead of once every 6 weeks).

4.2.5 Perspective, time horizon and discounting

The analysis was performed from a NHS and Personal Social Services (PSS) perspective, in line with the NICE reference case.[77] The model used a lifetime time horizon, following a patient until either death or an age of 100 years, which was in line with both the NICE reference case and European Society for Clinical and Economic Aspects of Osteoporosis (ESCEO)/ International Osteoporosis Foundation (IOF) guidelines.[78] All costs and benefits, i.e., life years and QALYs gained, were discounted at an annual rate of 3.5% as per the NICE reference case.[1]

4.2.6 Treatment effectiveness and extrapolation

4.2.6.1 Fracture incidence

General population risk of fractures

The model inputs for the general population risk of hip, vertebral and non-hip, non-vertebral (NHNV) fractures were the same as those estimated using the method described in the IOF/ European Federation of Pharmaceutical Industry Associations (EFPIA)-endorsed study on osteoporosis in the European Union by Hernlund et al. 2013 and reported for women in various age categories from the UK in the accompanying compendium of country-specific reports by Svedbom et al. 2013.[79, 80] The incidence of hip fractures were sourced from a study by Singer et al. 1998, which was considered as the most comprehensive data on hip fracture incidence in the UK.[81] According to the company, the study by Singer reported similar findings to a more recent UK study using the Clinical Practice Research data (CPRD) link over the years 1990-2012 (i.e., van der Velde et al. 2016 which also showed that the incidence of hip fractures remained stable over the studied time period.[82] Due to unavailability of data

on the risk of clinical vertebral fractures in the UK, the incidence of vertebral fractures was estimated based on the ratio of clinical vertebral to hip fractures in a Swedish study.[83] The incidence of NHNV fractures was estimated based on a combination of the incidence of forearm fractures (distal forearm, distal radius and wrist) that was sourced from Singer et al. 1998,[81] and the ratio of "other fractures" (femur, pelvis, humerus, rib, clavicle, scapula and sternum) to hip fractures in Sweden applied to the incidence of hip fractures as estimated by Singer et al. 1998 for the UK.[80, 83] The selected inputs for incidences of hip, vertebral and NHNV fractures are displayed in Table 4.6.

Table 4.6: Incidence of fracture per 100,000 people in the UK by age

Age	Hip[81]	Vertebral[83]	NHNV[81, 84]
50–54	33	84	633
55–59	51	142	813
60–64	81	143	979
65–69	132	192	1,425
70–74	282	397	1,928
75–79	619	602	2,891
80–84	1,236	777	3,876
85+	2,255	1,061	5,958

Based on Table 18 of the CS[1]
CS = company submission; NHNV = non-hip, non-vertebral

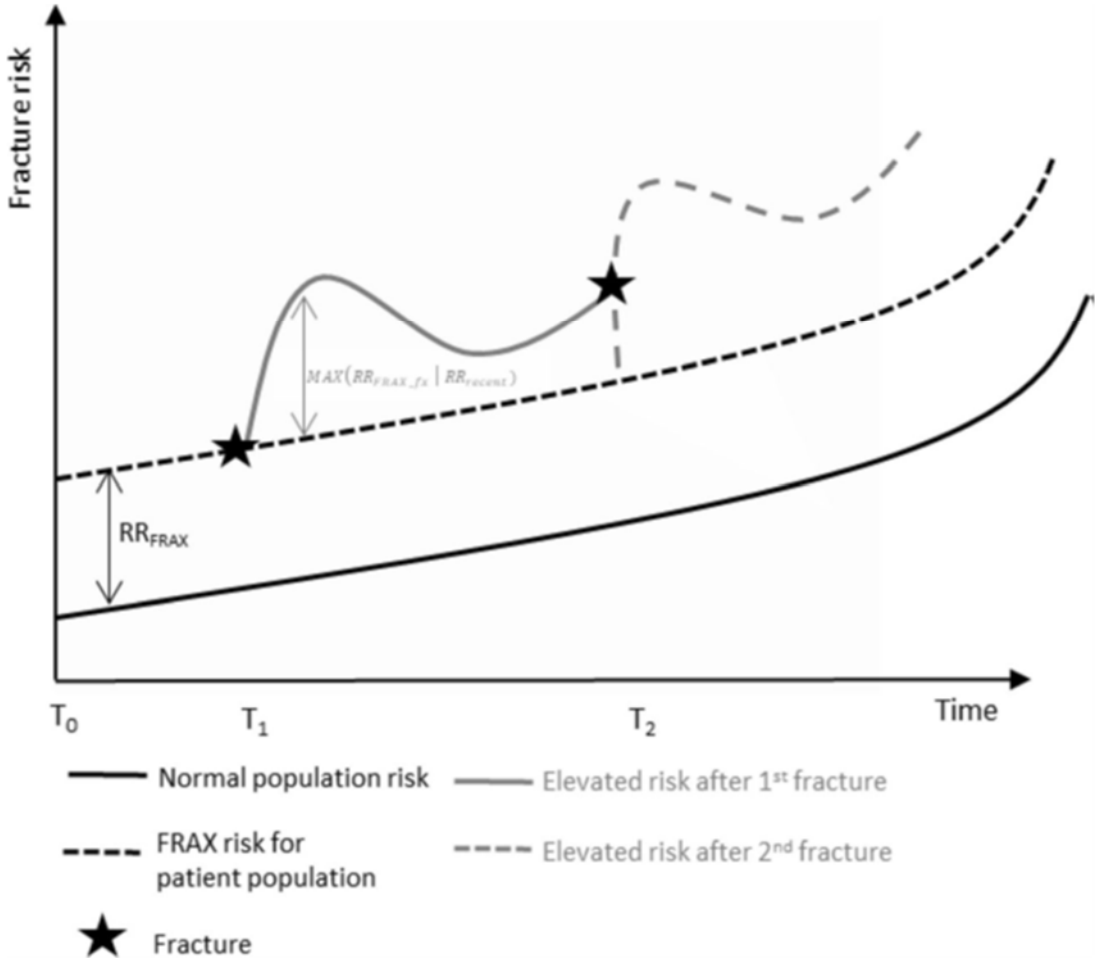
Increased fracture risk associated with osteoporosis

The model inputs for the increased fracture risk associated with osteoporosis, relative to the general population, were based on the FRAX algorithm. The FRAX tool, similar to QFracture, can be used to estimate an average 10-year risk of fracture based on clinical risk factors including age, BMI, BMD and lifestyle factors (e.g., smoking). The use of fracture risk assessment tools, such as FRAX and QFracture, in clinical practice is recommended by NICE clinical guideline (CG) 146.[11] The company preferred to use FRAX over QFracture because FRAX can be used in combination with BMD, is more widely used than QFracture, is included in the National Osteoporosis Guideline Group (NOGG) 2017 clinical guideline,[76] and can be more easily adapted to also consider the imminent fracture risk.

Imminent fracture risk

The model inputs for the imminent fracture risk, defined as the increased risk of a subsequent fracture after having sustained a first, second or third fracture, were sourced from Söreskog et al. 2020.[85] This study made use of a large dataset obtained from a retrospective real-world study in Swedish women aged 50 years and over with a fragility fracture[86], and estimated HRs for the risk of MOF in women after one, two or three fractures, relative to age- and gender-matched controls. The imminent fracture risk reaches its peak level in the first year following a fracture and then slowly declines until there is little excess risk after 5 years. When subsequent fractures occur within the timeframe of imminent risk following a prior fracture, the increases in risk may accumulate over time as "fracture cascades". An illustration of an individual patient's risk trajectory is shown in Figure 4.2 for a patient without a fracture at baseline. In contrast, the company's base-case model does assume a recent fracture at baseline.

Figure 4.2: Illustration of the fracture risk trajectory estimated using imminent risk



Based on Figure 14 in the CS,[1] which was sourced from Söreskog et al. 2020.[85]
Note: In contrast to the illustration above, the company's base-case model does assume a recent fracture at baseline.
MAX = maximum; RR_{FRAX} = relative risk estimated by FRAX for a given patient profile excluding prior fracture as a clinical risk factor; RR_{FRAX_fx} = relative risk estimated by FRAX for a given patient profile including prior fracture as a clinical risk factor; RR_{recent} = relative risk of an imminent fracture; T_0 = timepoint 0, at which the patient has no fracture history; T_1 = timepoint 1, at which the patient has sustained the first fracture; T_2 = timepoint 2, at which the patient sustained the second fracture.

Total fracture risk

For patients in the model, fracture risk was calculated as a function of the UK general population risk, the RR from FRAX for a given patient profile excluding prior fracture as a clinical risk factor, the maximum of the RR due to a recent fracture vs. no fracture (i.e. the imminent risk) or the RR from FRAX for a given patient profile including prior fracture as a clinical risk factor vs. the general population, and the risk reduction from treatment (see Section 4.2.6 of the ERG report). The formula that was used for this calculation is the following:

$$MAX(RR_{FRAX\_fx} \mid RR_{recent}) * RR_{FRAX} * General\ population\ risk * Risk\ reduction\ from\ treatment,$$

where MAX = maximum; RR_{FRAX} = relative risk estimated by FRAX for a given patient profile excluding prior fracture as a clinical risk factor; RR_{FRAX_fx} = relative risk estimated by FRAX for a given

patient profile including prior fracture as a clinical risk factor; RR_{recent} = relative risk of an imminent fracture.

Reduction of fracture risk

Efficacy estimates for romosozumab/alendronate and the comparators were applied to the above baseline fracture risks. The base-case efficacy estimates for romosozumab vs. alendronate were determined from the fracture endpoints from the ARCH study.[3] In analyses vs. other comparators, efficacy was estimated using an NMA. Treatment effects were estimated on the trial ITT population.

ARCH was considered the most relevant source of clinical evidence for modelling patients at imminent risk of fracture as it is the only study of romosozumab in women with prior fracture which includes fracture outcomes. Time-to-event analysis of fracture incidences are available from the clinical study report (CSR) for clinical fracture, non-vertebral fracture, hip fracture, and MOF. Cumulative point estimates are published for 12 and 24 months for new vertebral, clinical, non-vertebral and hip fracture types.[3]

Time-dependent efficacy of romosozumab/alendronate vs. alendronate alone were calculated for hip and non-vertebral fracture for each six-months cycle based on a continuous hazards approach using data from ARCH.[1] Patient-level data for each treatment arm was reconstructed from the published Kaplan-Meier curves. Parametric distributions were fitted to the model, and time-dependent hazard rates were calculated for the mid-point of the model cycle. In the model, efficacy of non-vertebral fractures was applied to NHNV fractures due to lack of data on all fractures excluding both hip and vertebral. For vertebral fractures, efficacy of new vertebral fractures was calculated from the published data at 12 and 24 months.[3] Efficacy for vertebral fractures beyond month 24 is based on 24 month efficacy. The resulting non-cumulative HRs of romosozumab/alendronate vs. alendronate are displayed in Table 4.7.

Table 4.7: ARCH non-cumulative efficacy data based on parametric distributions. HR of romosozumab/alendronate vs. alendronate by time point. ITT population.

Time since treatment start (months)	HR (hip fracture)	HR (new vertebral fracture, used for vertebral fracture in the model)	HR (non-vertebral fracture, used for NHNV fracture in the model)
0–6	[redacted]	[redacted]	[redacted]
7–12	[redacted]	[redacted]	[redacted]
13–18	[redacted]	[redacted]	[redacted]
19–24	[redacted]	[redacted]	[redacted]
25–30	[redacted]	[redacted]	[redacted]
31–36	[redacted]	[redacted]	[redacted]

Based on Table 19 of the CS[1]
CS = company submission; HR = hazard ratio; ITT = intention-to-treat; NHNV = non-hip, non-vertebral

ARCH compared romosozumab/alendronate to alendronate. Therefore, ARCH provides no efficacy data vs. placebo. In the model, fracture risk reductions from treatment are applied to the general population risk. Therefore, it was necessary to transform the ARCH efficacy of romosozumab vs. alendronate to romosozumab vs. placebo. To calculate RRs for romosozumab/alendronate vs. no treatment, the HRs of romosozumab/alendronate vs. alendronate alone in Table 4.7 above were applied to RRs of alendronate vs. placebo derived from the NMA (described in Section 3.4 and below). Since HRs (Table 4.7) and RRs (from the NMA) were similar, the company assumed, given the lack of RR

data from ARCH, that these could be used interchangeably.[1] The approach of using the alendronate vs. placebo data was considered reasonable given that, according to the company, the efficacy data of alendronate vs. placebo from the CS NMA do not differ significantly from other NMAs, for example NICE's most recent NMA. A comparison of results from the NMA in the current submission compared to the NMA from NICE's most recent NMA is provided in Table 4.8 below.

Table 4.8: Comparison of results in the NMA included in this submission to the most recent NICE Assessment Group NMA

Time since treatment	Time since treatment start (months)	Hip fracture (CS ITT NMA)	Hip fracture (NICE AG NMA)[a]	Vertebral fracture (CS ITT NMA)	Vertebral fracture (NICE AG NMA)[a]	Other (NHNV) (CS ITT NMA)	Other (NHNV) (NICE AG NMA)[a]
Romosozumab/ alendronate vs. placebo	0–12	[redacted]	0.39 (0.21 to 0.72)	[redacted]	0.25 (0.15 to 0.43)	[redacted]	0.71 (0.48 to 0.85)
	13–24	[redacted]		[redacted]		[redacted]	
	25–60	[redacted]		[redacted]		[redacted]	
Alendronate vs. placebo	0–12	[redacted]	0.64 (0.45 to 0.88)	[redacted]	0.50 (0.40 to 0.64)	[redacted]	0.77 (0.64 to 0.90)
	13–24	[redacted]		[redacted]		[redacted]	
	25–60	[redacted]		[redacted]		[redacted]	
Teriparatide vs. placebo[b]	0–12	[redacted]	0.35 (0.15 to 0.73)	[redacted]	0.23 (0.16 to 0.32)	[redacted]	0.58 (0.45 to 0.76)
	13–24	[redacted]		[redacted]		[redacted]	

Based on Table 20 of the CS[1]

[a] RRs in the NICE NMA were not calculated at specific timepoints; [b] Twelve-months efficacy for hip fracture was not available for the respective comparison with teriparatide in the CS NMA; twenty-four months efficacy was therefore assumed for the first 24 months for these treatments.

AG = assessment group; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis

The NMA provided efficacy estimates up to 36 months from treatment initiation. The treatments with longer treatment durations, efficacy is extrapolated beyond 36 months until the end of the treatment duration, in line with the independent academic Assessment Group's approach in the suspended NICE multiple technology appraisal (MTA) ID901.[87] Table 4.9 presents the base-case efficacy input of romosozumab/alendronate vs. placebo, where efficacy has been calculated based on the NMA using the ITT population. A scenario analysis was also conducted using the EU-label matched NMA (described in Section 3.4 based on the results presented in Appendix D.4.4).[8] The corresponding efficacy inputs for romosozumab/alendronate vs. placebo are presented in Table 4.10.

Table 4.9: Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and NMA (ITT populations)

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	NHNV fracture
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Based on Table 21 of the CS[1] and Table 48 of the response to request for clarification.[9]
CI = confidence interval; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis.

Table 4.10: Fracture risk ratio (95% CI), by fracture type and time point of romosozumab-to-alendronate vs. placebo based on the ARCH trial and scenario NMA (EU label-matched population)

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
Romosozumab-to alendronate vs. placebo (ARCH/ NMA)	0–6	[redacted]	[redacted]	[redacted]
	7–12	[redacted]	[redacted]	[redacted]
	13–18	[redacted]	[redacted]	[redacted]
	19–24	[redacted]	[redacted]	[redacted]
	25–30	[redacted]	[redacted]	[redacted]
	31–36	[redacted]	[redacted]	[redacted]
	37–42	[redacted]	[redacted]	[redacted]
	43–48	[redacted]	[redacted]	[redacted]
	49–54	[redacted]	[redacted]	[redacted]
	55–60	[redacted]	[redacted]	[redacted]

Based on Table 22 of the CS[1]

Drug	Time since treatment start (months)	Hip fracture	Vertebral	NHNV
CI = confidence interval; CS = company submission; EU = European Union; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis				

The NMA described in Section 3.4 was used to conduct scenario analyses for romosozumab/alendronate vs. other comparators, including teriparatide, denosumab, zoledronate, risedronate and raloxifene.[1] These scenarios were based on the NMA using the ARCH and FRAME ITT population, presented in Table 4.11. For completeness, the equivalent analysis performed using the EU label-matched NMA is presented in Table 24 of the CS (data only available 12-monthly instead of 6-monthly in the base-case NMA).[1]

Table 4.11: Fracture risk ratio (95% CI), by fracture type, based on network meta-analysis (NMA, ARCH and FRAME ITT population)

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	Other fracture (NHNV)
Romosozumab/ alendronate vs. placebo	**0–12[a]**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Alendronate vs. placebo	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Teriparatide vs. placebo[b]	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
Denosumab vs. placebo	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Zoledronate vs. placebo	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Risedronate vs. placebo[b]	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]
Raloxifene vs. placebo	**0–12**	[redacted]	[redacted]	[redacted]
	13–24	[redacted]	[redacted]	[redacted]
	25–60	[redacted]	[redacted]	[redacted]

Based on Table 23 of the CS[1]

[a] Results from FRAME are only included at month 12; results for romosozumab/alendronate from month 13 onwards only include ARCH, as discussed in Sections 3.2.7 and 3.3.2; [b] Twelve-months efficacy for hip fracture was not available for the respective comparison with teriparatide and risedronate. Twenty-four months efficacy was therefore assumed for the first 24 months for these treatments.

Drug	Time since treatment start (months)	Hip fracture	Vertebral fracture	Other fracture (NHNV)
CI = confidence interval; CS = company submission; ITT = intention-to-treat; NHNV = non-hip, non-vertebral; NMA = network meta-analysis.				

Fixed effect models were used for all fracture endpoints and time periods since the deviance information criterion (DIC) was lower in the fixed effect models compared with the random effect models, as shown in Appendix D.4.3 and D.4.4.[8]

As noted above, the results from the CS NMA do not differ significantly from other NMAs (Table 4.8) according to the company.[87, 88] However, one important difference is that the CS NMA considers time-specific results, unlike previously published NMAs, which have instead assumed equal efficacy across timepoints and only considered the final efficacy time point reported in each RCT.[1] By considering fracture outcomes at specific timepoints, the CS NMA was able to consider the short and long-term comparative efficacy of each osteoporosis treatment more accurately, compared to previously published NMAs. The importance of conducting a timepoint specific NMA is illustrated throughout the NMA results presented in Section 3.4 and Appendix D.4.3 and D.4.4, where it can be seen that treatment rankings and pairwise comparisons regularly varied across different time points for the same fracture outcomes.[1, 8] This is particularly important when considering bone-building treatments, such as romosozumab, which reaches the optimal clinical performance in a relatively short duration (i.e., 12 months), providing a rapid and potent effect and demonstrating the potential to interrupt such a "fracture cascade" early in the process. The accurate consideration of short-term comparative efficacy (i.e. at Month 12) is of particularly importance for patients who have incurred a recent MOF within the past 24 months and are at imminent risk of another fragility fracture, as these patients will experience particular benefit from osteoporosis treatments with fast-acting benefits.[89]

ERG comment: During the clarification phase, the ERG requested the company to explain the extent to which fracture incidences in the UK have remained stable over time and similar to those in the Singer et al. 1998 study.[81] The company responded by referring to the study by van der Velde et al. 2016 that made use of the CPRD data from the years 1990 – 2012.[82]

In van der Velde et al. 2016, the incidence of hip fractures overall remained stable at about 35/10,000 person-years, or at about 50/10,000 person-years for women aged 75 to 79 years.[82] For women in the same age group in Singer et al. 1998 this incidence was 70.74/10,000 person-years (for women aged 70 to 74 years it was 48.5 and for women aged 80 to 84 years it was 143.72). The ERG concludes that the study by van der Velde et al. 2016 indeed confirms the stability of fracture incidence over time, but also that the incidence rates in this study are substantially lower than in the study by Singer et al. 1998.[81] As such, the validity of the incidences of hip fractures that are used in the model is uncertain.

During the clarification phase, the company explained that they had not used the estimates from Singer et al. 1998 for clinical vertebral fractures because these were deemed unrealistically low in comparison to other studies. The company indicated that that could be due to vertebral fractures being treated in other healthcare facilities than those that were included in the study. Therefore, the company estimated incidence of vertebral fractures based on the ratio of clinical vertebral to hip fractures in a Swedish study.[83] The company clarified the validity of this ratio for the UK by referring to a study by Kanis et al. 2001 that, according to the company, showed that these ratios are similar between Sweden and the UK.[90] However, the ERG notes that Kanis et al. 2001 did not include an actual comparison between the

ratios of clinical vertebral to hip fractures in Sweden and the UK. As such, the validity of the incidences of vertebral fractures that are used in the model is uncertain. In response to a request by the ERG, the company performed a scenario analysis using estimates of vertebral fracture incidence by Singer et al. 1998 that resulted in an ICER that was almost twice the value of the company's base-case ICER.

The company indicated that although the incidence of radius/ulna fractures in the UK has decreased in the year 1998 relative to preceding years, it remained stable in the years 1998 – 2012 at approximately 40/10,000 person-years in the van der Velde study. Regarding the extent of similarity for the incidences of forearm fracture between the studies by Singer et al. 1998 and van der Velde et al. 2016, the company indicated that the incidences of wrist fracture in women aged 75-79 was approximately 70/10,000 person-years in Singer et al. 1998 and approximately 50-70 per 10,000 persons-years in van der Velde et al. 2016. The ERG notes that the latter incidence refers to distal forearm fractures, which is a combination of fractures in the radius/ulna and wrist (i.e., carpal fractures). The ERG notes that in the study by van der Velde et al. 2016 the incidence of wrist fracture was stable in the years up to 1998, but has doubled in the time period 1998 to 2012 and that in the study by Singer et al. 1998 the incidence of forearm fractures (i.e., radius / ulna) was 0.68 / 10,000 person- years in women aged 75-79 years. The ERG therefore concludes that the incidence has indeed remained stable over time for radius/ulna fractures but not for wrist fractures, and that the similarity of the estimates for forearm fracture incidence is low between the two studies. The company did not comment on the similarity between the ratios of the incidence of "other fractures" relative to hip fractures in Sweden and the UK. As such, the validity of the incidences of NHNV fractures that are used in the model is uncertain.

The model uses relative risk values for the imminent fracture risk that were sourced from the study by Söreskog et al. 2020.[85] It is not clear to the ERG how the values that are used in the model correspond to those reported by Söreskog et al. 2020, which is possibly due to the use of different age categories in the paper and the model. The model also specifies values of 0 for the relative risk of a 4th fracture after a 3rd fracture, in contrast to Söreskog et al. who report non-zero values for this. No explanation was provided for this aspect; therefore, it is not clear to the ERG what the underlying rationale is for the assumed 0 values.

The 'State trace' sheet of the model provides an overview of the proportions of patients having sustained their 1st, 2nd, 3rd, and 4+ hip, vertebral or NHNV fractures. Logically, over time first a proportion of patients has their first fracture, followed by a second, et cetera. However, the proportion of patients that has their first NHNV fracture remains zero throughout the model time horizon whilst there is a non-zero proportion of patients having their second NHNV fracture from the second cycle of the model onwards. The ERG could not trace the root cause of this inconsistency.

The company has assumed that the relative risks of fracture after having had a 1st, 2nd or 3rd fracture as estimated using Swedish data are transferable to the UK. To support this assumption during the clarification phase, the company referred to the geographical proximity and similarity in quality of healthcare between Sweden and the UK and the fact that previous CE studies have made the same assumption. According to the ERG, the validity of the assumption that the relative risks of fracture are transferable between the two countries is not sufficiently justified.

In previous publications based on the same model by Söreskog et al. 2020 a limitation was noted in relation to the imminent fracture risk being possibly overestimated, because not all risk factors that are included in FRAX were available to adjust the imminent risk ratios for confounding.

To conclude, the ERG is uncertain regarding the validity of the values used for the imminent fracture risk as well as regarding their implementation in the model. In response to a request by the ERG, the

company performed a scenario analysis using only the FRAX algorithm, which includes a risk factor for prior fracture regardless of fracture recency, that resulted in an increase in the ICER, becoming more than twice the company's base-case ICER.

Treatment effect on fracture risk of romosozumab/alendronate vs. alendronate alone was calculated by reconstructing patient-level data from published Kaplan-Meier curves and then fitting parametric distributions in order to calculate time-dependent hazard rates. These (survival data) analyses are not shown in the CS. In response to clarification question B7.B, the company mentioned that the analyses were conducted internally but they are not publicly available.[9] While the methods used for the survival analyses seem appropriate, it should be emphasised that the results of such analyses were are not presented. Therefore, the ERG cannot assess whether the distributions were properly fitted and cannot explore the impact of using alternative distributions on the model results.

4.2.6.2 Persistence

Suboptimal persistence to osteoporosis medications is frequently observed in UK clinical practice, and may reduce the treatment efficacy and increase the risk of fracture compared to the reduction in fracture risk seen with optimal persistence.[1] One UK-based study (N=63,350) found that 50% of all women receiving osteoporosis treatments had discontinued treatment after six months, with 68% of all women discontinuing by the end of one year.[91]

To account for this in the model all patients were at risk of treatment discontinuation in each cycle, with discontinuation reflected in their anti-fracture treatment benefits. In the base-case, patients were assumed to be at risk of discontinuation during the first three years, after which persistence remained stable until treatment was completed, based on long-term studies indicating that discontinuation rates are highest immediately after the initiation of treatment, with discontinuation rates plateauing and remaining stable after the first year and up to five years of treatment.[92, 93] A treatment duration of five years was assumed to align with previous health economic studies and recommendations from ESCEO/IOF.[78, 94, 95]

Patients who discontinued treatment could not switch to, or restart, a treatment, due to the lack of sequential evidence in the published literature, as most RCTs have been conducted in treatment naïve patients, or required a long treatment washout period prior to enrolment.[1] For persistent patients who switch treatment within a sequence in the model, patients were assigned the probability of non-persistence corresponding to the time since the start of the treatment.

In the base-case, persistence on alendronate alone was derived from Li et al. 2012, who used the UK General Practice Research Database (GPRD) to estimate persistence on osteoporosis medications among postmenopausal women in the UK.[96] In scenarios, persistence on risedronate and raloxifene were also estimated from Li et al. 2012. Persistence on denosumab was taken from a retrospective observational study using the Swedish Prescribed Drug Register,[97] while persistence on teriparatide and zoledronate were taken from a Swedish osteoporosis database.[75]

Persistence on romosozumab in clinical practice is unknown. As a starting point the company considered the persistence on teriparatide. A Swedish osteoporosis database reported that teriparatide had a 6-month and 12-month persistence of approximately 74% and 61%, respectively.[75] The company argue that as romosozumab will be administered much less frequently compared to teriparatide (QM vs QD), and UCB will provide a PSP in the UK, it is reasonable to assume that persistence on romosozumab will be higher than on teriparatide.[1] However, the size of this improvement is unknown. Based on the three pivotal romosozumab clinical trials,[3, 17, 18] the company assumed that 90% of patients

will be persistent to treatment throughout the 12-month romosozumab treatment period. In ARCH, [redacted]% of patients receiving romosozumab completed the first 12-month treatment period.

For the treatment sequence of romosozumab followed by alendronate used in this submission, it was assumed that the persistence rates for alendronate would be 85% of the persistence for denosumab. This was based on the assumption that patients who initially demonstrated high persistence on romosozumab would be expected to demonstrate high persistence on follow-on treatments, and therefore the persistence on alendronate after romosozumab would be notably higher than the persistence on alendronate alone reported by Li et al. 2012.[96] The company report that this assumption is supported by a study of persistence to treatment in chronic diseases, which found that patients who have already persisted on treatment for a year have a 50% reduced discontinuation rate compared to patients just starting treatment.[98] Additionally the company note that the patient population in Li et al. 2012 is less severe that the target population for romosozumab, as they were not required to have experienced a previous fracture, while patients eligible for treatment with romosozumab/alendronate will have experienced a recent MOF within 24 months.[96] The company would expect that these more severe patients would exhibit improved persistence and that USB's PSP will include support with the transition to follow-on treatment, which is likely to further increase persistence on alendronate after romosozumab compared to alendronate alone.[1] A summary of persistence assumptions for all treatments can be found in Table 4.12.

Table 4.12: Proportion of patients on osteoporosis treatment over time in the economic model

Month since treatment initiation	Romosozumab	Alendronate after romosozumab[a]	Alendronate alone[96]	Teriparatide[b75]	Zoledronate[b75]	Denosumab[b97]	Risedronate[b96]	Raloxifene[b96]
6	90%	85%	49%	74%	100%	100%	50%	45%
12	90%	71%	38%	61%	100%	83%	38%	33%
18	0%	59%	34%	3%	51%	69%	33%	30%
24	0%	53%	30%	3%	42%	62%	28%	26%
30	0%	47%	27%	0%	34%	56%	24%	23%
36	0%	43%	24%	0%	28%	50%	21%	21%
42	0%	38%	22%	0%	23%	45%	18%	19%
48	0%	34%	20%	0%	18%	40%	16%	17%
54	0%	31%	19%	0%	15%	36%	14%	16%
60	0%	28%	17%	0%	12%	33%	12%	14%

Based on Table 25 of the CS[1]
[a] The persistence on alendronate after romosozumab was assumed to be 85% of the persistence on denosumab;
[b] Treatment included in scenario analyses only.
CS = company submission

ERG comment: The company's approach to model persistence is inconsistent between intervention (romosozumab) and comparators and is likely to be biased in favour of romosozumab. The guidelines for economic evaluations in osteoporosis endorsed by the ESCEO/IOF recommend using real-world data on medication adherence.[99] However, this approach was only used for the comparators.

The company assumed that persistence with romosozumab is 90%, which was based on persistence with romosozumab as observed in the ARCH trial. However, in response to clarification question B9 the company indicated that "*persistence data from retrospective observational studies are more appropriate than persistence data from clinical trials. Persistence in clinical trials is significantly higher than in clinical practice most likely because patients know they are being observed and have consented to participate in the study*" and that "*persistence of romosozumab is assumed to be the same as in the ARCH trial, despite clinical trials show higher persistence than what is seen in clinical practice. This was necessary given that there is no real-world evidence currently available for romosozumab as it has only been recently launched*".[9] The ERG agrees that real-world persistence with romosozumab, outside the context of a clinical trial, will be lower than in ARCH and therefore prefers to use a lower value for their base-case analysis. In line with the assumption made by Söreskog et al. 2021 in their CE analysis for romosozumab in Sweden,[73] the ERG assumes a value of 80% for persistence with romosozumab. The ERG considers this a plausible value since it is lower than persistence with romosozumab in ARCH and higher than the real-world persistence with teriparatide that the company sourced from the Swedish osteoporosis database. The latter is supported by the notion that romosozumab will be administered less frequently than teriparatide and that it is likely that persistence with romosozumab is higher relative to treatments with higher frequencies of administration.

For persistence with alendronate, the company assumed lower values for persistence with alendronate alone than for persistence with alendronate after romosozumab. Specifically, the company assumed that persistence with alendronate after romosozumab is 85% of persistence with denosumab as sourced from a Swedish study by Karlsson et al. 2015.[97] The ERG considers this an arbitrary choice. The company sourced persistence with alendronate alone from Li et al. 2012,[96] which was a study on persistence with osteoporosis therapies based on UK CPRD data. The company justified the use of different sources by referring to a difference in the severity of osteoporosis between patients treated with either alendronate after romosozumab or alendronate alone. Since alendronate, as a standalone treatment, was positioned as the most relevant comparator to romosozumab in the indicated population for the company's base-case analysis, the ERG considers it inappropriate to assume a difference in severity of osteoporosis for the population that is considered eligible for both treatment options. Therefore, the ERG prefers to inform persistence with alendronate, regardless of whether it is given as a standalone treatment or after romosozumab, using the same study. Furthermore, the ERG was unable to verify the persistence values shown in Table 4.12 that the company indicated were sourced from Li et al. 2012.[96] Importantly, the data in the study by Li et al. 2012 range from 1995 to 2008 and indicate that persistence estimates have not been stable over that period of time. The ERG identified a more recent study by Morley et al. 2020 on persistence with osteoporosis therapies that also made use of UK CPRD data.[100] The ERG preferred to use this more recent source of persistence estimates for their base-case, and used the estimates for non-naïve patients for alendronate after romosozumab and the estimates from naïve patients for alendronate alone.

In addition to persistence with alendronate after romosozumab and alendronate alone, the ERG also used the study by Morley et al. 2020 to inform persistence with denosumab, risedronate and raloxifene using data from the subgroup of naïve patients.[100] Whilst Morley et al. 2020 also provide estimates for persistence with teriparatide and zoledronate, the ERG did not use these estimates because they were based on very small (n<20) sample sizes. Instead, the ERG preferred to use the same estimates as the company for persistence with these comparators. However, the ERG did not have access to the Swedish osteoporosis database that informed these estimates nor any details regarding the methods that were

used. As such, the validity of these estimates remains uncertain. The ERG preferred estimates of persistence are presented in Table 4.13.

Table 4.13: ERG preferred estimates of persistence with osteoporosis therapies

Month since treatment initiation	Romosozumab[73]	Alendronate after romosozumab[a100]	Alendronate alone[b100]	Teriparatide[c,d]	Zoledronate[c,d]	Denosumab[c100]	Risedronate[c100]	Raloxifene[c100]
6	80%	31%	62%	74%	100%	64%	62%	53%
12	80%	19%	51%	61%	100%	55%	51%	42%
18	0%	14%	44%	50%	51%	48%	44%	37%
24	0%	11%	38%	41%	42%	36%	38%	33%
30	0%	9%	34%	0%	34%	32%	34%	29%
36	0%	8%	29%	0%	28%	28%	29%	25%
42	0%	7%	26%	0%	23%	25%	26%	24%
48	0%	6%	24%	0%	18%	22%	24%	24%
54	0%	5%	21%	0%	15%	19%	21%	23%
60	0%	4%	18%	0%	12%	16%[c]	18%	22%

[a] Based on estimates from non-naïve patients in Morley et al. 2020; [b] Based on estimates from naïve patients in Morley et al. 2020; [c] Treatment included in scenario analyses only; [d] Same values as company base-case
CPRD = Clinical Practice Research data; ERG = Evidence Review Group; PSP = patient support programme; UK = United Kingdom

The company indicated in the CS that differences in persistence exist between patients that previously persisted on osteoporosis treatment (i.e., non-naïve patients) and patients that just started with osteoporosis treatment (i.e., naïve patients). For example, Morley et al. 2020 found that persistence with oral bisphosphonates was higher in naïve patients than in non-naïve patients.[100] This contrasts with findings from an earlier study that found the opposite.[101] Also, the company considered that the PSP is likely to increase persistence on alendronate after romosozumab compared to alendronate alone. However, this assumption is not based on any evidence. To address the uncertainty surrounding this aspect and the extent to which patients can still be considered as naïve once they have persisted with a six-month treatment course, the ERG assessed the impact on the CE results when assuming the same persistence for naïve and non-naïve patients. For this the ERG performed a scenario analysis in which persistence was based on the pooled data from all patients (i.e., both naïve and non-naïve patients) in Morley et al. 2020, for both alendronate alone and for alendronate after romosozumab.[100]

4.2.6.3 Dynamic residual effects

The company assume that the time a patient remains on osteoporosis treatment is directly related to the duration of efficacy that can be expected. They argue that there is consensus that anti-fracture efficacy

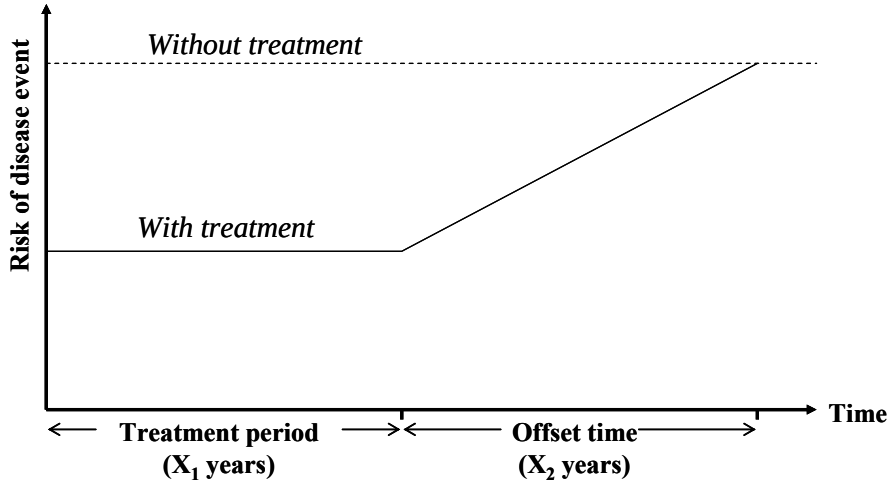
persists for a period of time (offset time) after treatment is discontinued in patients with osteoporosis.[102] Two alternatives for modelling residual effects are presented in the CS and in Figure 4.3 below:[1]

- Dynamic: Offset time is assumed to be as long as time on treatment and is, therefore, shorter for patients who drop out earlier. Partially persistent patients are distributed over a range of treatment durations and corresponding offset times depending on if and when they stopped treatment.
- Fixed: All patients have the same specified offset time irrespective of treatment drop out, so a patient who discontinues after 1 year will nonetheless have 2-years offset time if the prespecified offset time was 2 years.

During the offset time the fracture risk reduction is assumed to decline linearly to zero.[1] The efficacy of the last treatment given to the patient in the sequence was used for the offset time. Thus, if a patient was treated with romosozumab for 12 months and alendronate for the following 36 months, the offset time equalled 48 months and efficacy used for offset was based on the efficacy of alendronate for patients who had previously received romosozumab. This was validated by leading UK experts at an advisory board. This approach is recommended by the ESCEO and IOF guidelines, and has been used in other published health economic studies and romosozumab HTA submissions to the Scottish Medicines Consortium (SMC) and TLV (Tandvårds- och läkemedelsförmånsverket, The Swedish Dental and Pharmaceutical Benefits Agency).[78, 103, 104]

The company report that evidence supports the assumption that alendronate, zoledronate and teriparatide have offset times similar to the treatment length and there is no robust evidence to support differential offsets for other treatments, providing evidence for the dynamic model approach.[105-109] For denosumab, efficacy was limited to 6 months after discontinuation.[110, 111] Chronic treatment with denosumab is necessary when used as the subsequent treatment after romosozumab for this combination to provide optimal benefits to patients; or alternatively a further treatment switch to a bisphosphonate after the denosumab treatment period would be required. In the model, a one-year fixed offset time was applied to denosumab.[1] This was described by the company as a conservative approach. A summary of the treatment sequences and associated length of effects is presented in Table 4.14 (a complete description of the scenarios is given in Section 5.2.3).

Figure 4.3: Modelling the residual effects of osteoporosis treatments



Based on Figure 15 of the CS.[1]
CS = company submission; X_1 = treatment period; X_2 = offset time

ERG comment: The company assumptions regarding dynamic residual treatment effects are broadly in line with the recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019.[78] Therefore, the ERG considers the company's approach appropriate. Scenarios with fixed offset time can be deemed as exploratory.

As described in Key issue 2, a scenario analysis where treatment waning starts at four years followed by a dynamic offset (linear waning) of the treatment effect was explored by the ERG in Section 6.1.2. In this scenario it was assumed 4 years of full effect, a waning in effect for one more year (the waning assumption was to consider an effect between sequential alendronate and alendronate alone as assumed by the company) followed by a dynamic offset 5 years. Note, however, that the other scenario mentioned in Key issue 2, one with shorter duration of the dynamic offset of the treatment effect was not possible to run. In the model implementation, offset time is either dynamic and equal to the time on treatment, or fixed to 1 year. The rationale for the second scenario was that, if treatment effect waning is possible, the duration of the residual treatment effect might be less than the time on treatment. Thus, for the combination romosozumab/alendronate, the ERG wanted to explore a scenario where the offset time was three years instead of the five assumed in the model. The ERG was unable to run this scenario, which is expected to increase the ICER.

Finally, the ERG would like to note that residual effects for zoledronate could be longer than those assumed by the company.[112] However, the ERG was unable to change the model to incorporate this assumption. Cost effectiveness results including zoledronate as comparator might be underestimating the ICER.

Table 4.14: Summary of treatment sequences and treatment effect duration applied for the base-case and company scenario analyses

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Base-case scenario											
Intervention: ROMO/ALN	ROMO	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Comparator: ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 1											
ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 2											
TRP (24 months)	TRP	TRP	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect
Scenario 3											
TRP (18 months)	TRP	TRP (1/2) NONE (1/2)	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Dynamic offset	Dynamic offset	No effect	No effect	No effect	No effect	No effect	No effect	No effect

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Scenario 4											
TRP (biosimilar)/ALN	TRP	TRP (1/2) ALN (1/2)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 5											
TRP/ALN	TRP	TRP (1/2) ALN (1/2)	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 6											
RAL	RAL	RAL	RAL	RAL	RAL	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 7											
DEN	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect
Scenario 8											
RIS	RIS	RIS	RIS	RIS	RIS	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 9											
ZOL	ZOL	ZOL	ZOL	ZOL	ZOL	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year	Year 9	Year 10	Year 11+
Scenario 10											
ALN	ALN	ALN	ALN	ALN	ALN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	Dynamic offset	No effect
Scenario 11											
DEN	DEN	DEN	DEN	DEN	DEN	NONE	NONE	NONE	NONE	NONE	NONE
Tx. Effect*	Full	Full	Full	Full	Full	Fixed offset	No effect	No effect	No effect	No effect	No effect

Based on Table 43 of the response to request for clarification (question B1).[9]
*Treatment effect on fracture risk reduction.
ALN = alendronate, DEN = denosumab, RAL = raloxifene, RIS = risedronate, ROMO = romosozumab, TRP = teriparatide, Tx. = treatment, ZOL = zoledronate

4.2.6.4 Mortality

Mortality is captured in the model in three ways: age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor.[1] Age- and gender-specific mortality rates for the general population (all-cause mortality) in the UK were based on the years 2012–2014.[113] At the start of the model mortality risk is determined by UK general population all-cause mortality. When a patient sustains a fracture, the relative risk of death compared with the non-fractured population is applied to the normal population risk, and the relative risk was down-adjusted to 30% to adjust for higher frailty (i.e., increased risk of death due to other reasons than the fracture itself) in the fractured population.[1, 94, 114]

ERG comment: It is unclear why the company used UK Life Tables from 2012 to 2014.[113] In the ERG base-case, the most recent version (2017 to 2019) was used.[115]

Mortality related to hip and clinical vertebral fractures

For hip fractures, age-dependent relative risks of death were sourced from Jönssen et al. 2011,[103] a study on the CE of denosumab in Sweden. The estimated mortality during the first and subsequent years after hip fracture from a sample of 36,551 Swedish women with a main diagnosis of femur fracture between 1997 and 2001 were used to calculate standardised mortality ratios (SMRs) relative to the mortality of the Swedish age- and gender-matched general population in 2000. It was assumed that the SMRs based on Swedish data were generalisable to the UK. For vertebral fractures, the age-dependent relative risks of death were also sourced from Jönssen et al. 2011.[103] In that study, mortality was based on data from a Swedish sample that included 994 patients who sustained a clinical vertebral fracture in 1993 to 1994.[116] The age- and sex-dependent mortality was used to calculate SMRs in the same way as for hip fractures, but relative to the mortality of the Swedish general population in 1994. The relative risks of mortality compared to the normal population are presented in Table 4.15 below.

Table 4.15: Relative risk of mortality for hip and clinical vertebral fractures compared to the general population

Age	Hip fracture Year 1[103]	Clinical vertebral fracture Year 1[116]	Hip fracture Year 2+[103]	Clinical vertebral fracture Year 2+[116]
50 years	9.79	12.07	3.62	7.94
55 years	8.64	10.15	3.34	6.67
60 years	7.69	9.04	3.11	5.94
65 years	6.39	7.43	2.70	4.88
70 years	5.54	5.98	2.44	3.93
75 years	4.16	4.39	1.91	2.88
80 years	2.92	2.75	1.39	1.81
85 years	2.15	1.98	1.06	1.30
90 years	1.63	1.36	1.00	1.00

Based on Table 26 of the CS[1]
CS = company submission

Mortality relating to NHNV fractures

For NHNV fractures, the relative risk of death was calculated as a weighted average of the estimates of relative risks reported by Barret et al. 2003 using the proportions of different fracture types reported by Kanis et al. 2001.[90, 117] The company assumed that the relative risks of death after rib (30% of the included fractures) and clavicle/scapula/sternum (13% of the included fractures), which were not reported by Barret et al. 2003 were equal to one (i.e., no excess mortality). The same relative risk was used for all ages, which the company justified by referring to the variation in fracture distribution across age groups which was deemed to be small. The company notes that since the relative risk of death for NHNV fractures is known to increase with age,[116, 118, 119] the use of the same estimate for all age groups could lead to underestimation in younger and overestimation in older patients. The estimated mortality after NHNV fracture is shown in Table 4.16. It was assumed that women sustaining a fracture at NHNV sites were at increased risk of death only within the first year of fracture.

Table 4.16: Mortality during the first year following NHNV fractures

Fracture type	Fractures	Proportion	Relative risk of death
Rib	340	30%	1.0
Pelvis	47	4%	1.7
Proximal humerus	352	31%	1.4
Humeral shaft	117	10%	1.2
Clavicle, scapula, sternum[a]	145	13%	1.0
Other femoral	52	5%	1.8
Tibia, fibula	98	9%	1.1
All	1,151	100%	1.23

Based on Table 27 of the CS[1]
[a] No excess mortality reported, relative risk assumed to be equal to 1.0.
CS = company submission; NHNV = non-hip, non-vertebral

Comorbidity adjustment excess mortality

It has been reported that patients with osteoporosis have a higher degree of frailty compared to the general population and that excess mortality after a fragility fracture is not entirely attributable to the fracture event. A common assumption is that 30% of excess mortality is directly caused by the fragility fracture.[94, 114] Therefore, it was assumed that 30% of excess mortality after hip, clinical vertebral or NHNV fracture was associated with the fracture event.

The model also assumed that a patient would incur the highest risk of excess mortality, depending on previous fracture history. For example, if a patient sustained a hip fracture in cycle three and an NHNV fracture in cycle five, the excess mortality risk that was highest was incorporated (in this instance the second-year hip fracture excess mortality). The increased mortality was assumed to persist for 8 years, in line with the follow-up period in previous studies.[119, 120]

ERG comment: For the calculation of the relative risk of death for NHNV fractures, the company used the incidence of fractures for the age group of 65 to 69 years from Kanis et al. 2001.[90] The ERG notes that the incidence an older age group (e.g., 70 to 74 years or 75 to 79 years) would have made for a better match with the modelled population, but this is unlikely to have a substantial impact on the CE results.

The ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] suggest that only the excess mortality of hip and vertebral fractures should be included, as there is not yet enough evidence regarding NHNV fractures. However, there was a lack of consensus on this inclusion of excess mortality due to vertebral fractures amongst the 23 clinical and economic experts that were asked to review and validate the recommendations. In light of this, the ERG prefers to include excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNVfractures were also explored by the ERG in Section 6.1.2.

Modelling mortality with the FRAX algorithm

Some of the clinical risk factors that are inputted into the FRAX algorithm are known to contribute to mortality. Based on this, one of the outputs of the FRAX algorithm is the relative risk of pre-fracture mortality for the defined patient population.[1] This relative risk was used to adjust the baseline mortality of patients in the model, as well as mortality after fracture. However, this assumed that the pre-fracture relative risk of mortality obtained from FRAX did not change once a patient had experienced a fracture. This assumption was made as the relationship between clinical risk factors and mortality post-fragility fracture has not yet been investigated.[1]

Using mortality relative risks from the FRAX algorithm resulted in higher risk populations having a higher overall mortality (compared to lower risk populations), and thus benefiting less from avoiding fractures, compared to if the mortality adjustment was not included.[1]

The FRAX algorithm does not take into account other risk factors (not inputted into the FRAX algorithm) that may differentiate the mortality of osteoporosis patients compared to the general population. Consequently, the assumption that only a proportion of the excess mortality after fracture is related to the fracture event is made, as described above. The model uses the highest mortality in situations where both post-fracture mortality and FRAX-derived mortality need to be accounted for.

4.2.7 Adverse events

The company note several AEs that can be associated with osteoporosis regimens include upper gastrointestinal (GI) symptoms, osteonecrosis of the jaw (ONJ), hypocalcaemia, bone pain, atypical femoral fractures (AFFs), influenza-like symptoms, conjunctivitis, atrial fibrillation and stroke.[1] However, they report that due to lack of evidence, the model only includes gastrointestinal adverse events (GIAEs) that are associated with oral bisphosphonates, and excludes other AEs associated with osteoporosis, in line with other economic models and previous NICE appraisals of anti-osteoporotic treatments.[121, 122] The CS confirmed that no adjudicated events of ONJ or AFF were reported in the 12-month double-blind ARCH treatment phase.[3] During the open-label alendronate treatment phase, only one ONJ event occurred in each arm (<0.1% each in the alendronate/romosozumab and alendronate/alendronate arms) and six AFF events (two events (<0.1%) and four events (0.2%) respectively) were observed.[1]

An imbalance in serious adjudicated CV AEs was observed in the ARCH trial.[1] Romosozumab is therefore contraindicated for patients with previous myocardial infarction (MI) or stroke.[4] Given this contraindication, which was not an exclusion criterion in the ARCH trial, the company considered it reasonable to exclude CV AEs from the economic analysis. They stated that this approach aligned with the independent academic Assessment Group's approach in the suspended NICE MTA ID901.[87]

ERG comment: It was unclear whether all CV events in the ARCH trial occurred in individuals with a history of MI or stoke. If not, then the exclusion of those events which occurred in people who would not be contraindicated would be inappropriate. At clarification the ERG requested that the company

included CV AEs in the model according to the incidence in the ARCH trial.[123] In response, the company included a scenario utilising the relative risk of a CV-event based on the ARCH study, including only patients who do not have the contraindication of prior MI or stroke.[9] The post-hoc analysis of ARCH showed that patients randomised to romosozumab who did not have the contraindication (MI or stroke) at baseline, had a relative risk of major adverse CV events of [redacted] during the first [redacted] years after randomisation, compared with alendronate (subject incidence [redacted]% in romosozumab arm vs. [redacted]% in alendronate arm).[9] Costs and disutilities related to CV events are described in the relevant HRQoL and cost sections.

4.2.8 Health-related quality of life

4.2.8.1 Health state utility values

HRQoL was assessed in the ARCH trial at pre-determined time points, irrespective of fracture occurrence. The company considered it inappropriate to use this trial QoL data as it did not provide robust sensitive utility values for fracture health states.[1] The collected QoL data were also treatment specific, which the company expected would underestimate the potential QoL gain associated with treatment.

Therefore, the company preferred to use utility multipliers for fractures from the International Costs and Utilities Related to Osteoporotic Fractures Study (ICUROS) combined with UK general population values from Szende et al. 2014.[73, 74, 124] The ICUROS study was designed to assess the QoL impact of fractures on osteoporosis patients over time for use in CE modelling. It is the largest prospective study on osteoporosis quality by including over 7,000 patients in 12 countries, including 357 fractures experienced by patients in the UK.[73, 74] The ICUROS measured QoL using the EQ-5D as soon as possible after fracture occurrence regardless of treatment, and then at 4, 12 and 18 months after fracture, allowing the estimation of short- and long-term impact of osteoporotic fracture in real-world patients. ICUROS utilities were used by the independent Assessment Group in technology appraisal (TA) 464 and have also been used in economic evaluations of romosozumab for the TLV in Sweden and the SMC in Scotland.[11, 104, 125] The ESCEO/IOF also recommend using national ICUROS data if available or otherwise the international version. The utility multipliers for the first year after fracture and the second and following years are displayed in Table 4.17.

Table 4.17: Utility multipliers

Health state	Multiplier	Reference
First year after fracture		
Hip fracture	[redacted]	ICUROS
Vertebral fracture	[redacted]	ICUROS
Other NHNV fractures	[redacted]	ICUROS
Second and following years after fracture		
Hip fracture	[redacted]	ICUROS
Vertebral fracture	[redacted]	ICUROS
Other NHNV fractures	[redacted]	ICUROS

Based on Table 28 of the CS[1]
CS = company submission; ICUROS = International Costs and Utilities Related to Osteoporotic Fractures Study; NHNV = non-hip, non-vertebral

These multipliers were applied to the UK general population utility values estimated by Szende et al. 2014 shown in Table 4.18.[124] Disutilities for multiple fractures were applied in a multiplicative approach.

Table 4.18: UK General population utility values

Age	General population utility
50 years	0.849
55 years	0.804
60 years	0.804
65 years	0.785
70 years	0.785
75 years	0.734
80 years	0.734
Source: Table 29 of the CS[1] CS = company submission	

ERG comment: The ERG agrees with the approach of using fracture event utility multipliers from a large study rather than the ARCH data, which was collected at set times rather than on occurrence of fracture events. The ICUROS study included patient data from EuroQoL-5 Dimensions-3 Levels (EQ-5D-3L), time trade-off (TTO) and EuroQoL-Visual analogue scale (EQ-VAS). In the clarification response, the company clarified that multipliers were based on EQ-5D-3L data only, not the TTO or EQ-VAS data.[9] This aligns with the measurement aspect of the NICE reference case. In their clarification response the company also clarified that the utility multipliers obtained from the ICUROS study were based on data from all countries included in the study as UK specific multipliers are not currently available.[9] However, it would appear, given the similarity of the current multipliers with those used in ID901 (shown below) that the UK value set, which was used in ID901, was also used to estimate utility multipliers in this case.[87] Therefore, while utilities may be slightly affected by different reporting of health in different countries (for example due to different quality of treatment or interpretation of response options), utilities are not affected by different preferences across countries as the UK value set was used for all countries. This increases the likelihood that values are representative of UK utilities.

The multipliers included in this submission differ somewhat from those used in TA464 and ID901, as shown in Table 4.19.[11, 87] ID901 multipliers are fairly similar to those presented in this submission. However, the multipliers presented in TA464 suggest that hip and NHNV fractures have less impact on HRQoL compared to the current submission, while vertebral fractures have more impact. The company stated at clarification that the difference between the current submission and ID901 in NHNV fractures was due to the fact that UCB included more fracture types than ID901.[9] Detailed data from ICUROS on utilities for additional fracture types were found in the appendix of a study by Kanis et al. 2018.[126] Other differences with TA464 were considered to be due to the larger sample size available in the analysis by the company, which included around 3,000 fracture patients rather than just over 1,000 in the prior appraisal. These alternative sets of multipliers will be considered in a scenario to explore the sensitivity of results to multipliers used.

Table 4.19: Utility multipliers across submissions

Health state	ID3936	ID901	TA464
First year after fracture			
Hip fracture	[redacted]	0.55	0.69

Health state	ID3936	ID901	TA464
Vertebral fracture	[redacted]	0.68	0.57
Other NHNV fractures	[redacted]	0.805*	0.87**
Second and following years after fracture			
Hip fracture	[redacted]	0.86	0.85
Vertebral fracture	[redacted]	0.85	0.66
Other NHNV fractures	[redacted]	0.995*	0.99**

Based on CS[1], NICE TA464,[11], and AG report[87]

* ID901 provided multipliers for proximal humerus and wrist separately. The multipliers in the table above have been estimated as the mean of the proximal humerus and wrist values presented (year 1, 0.78+0.83/2 = 0.805; and year 2, 1.00+0.99/2 = 0.995); ** TA464 provided multipliers for shoulder and wrist separately. The multipliers in the table above have been estimated as the mean of the shoulder and wrist values presented. (year 1, 0.86+0.88/2 = 0.87; and year 2, 1.00+0.98/2=0.99)

AG = assessment group; CS = company submission; NHNV = non-hip, non-vertebral; NICE = National Institute for Health and Care Excellence; TA = technology appraisal

The multiplicative approach for accounting for the impact of multiple chronic or acute fractures has been used in previous appraisals.[11, 87] The way in which chronic multipliers were combined differs somewhat across appraisals. In TA464, if more than one fracture occurred then the chronic multipliers for each fracture was applied, but no more than one acute fracture was applied at any one time.[11] In their clarification response the company confirmed that they assumed that a maximum of two acute multipliers could be applied at once.[9] It is unclear which approach is more appropriate in this case, but the ERG could not test the impact of this assumption as changing the VBA code was not possible.

The ERG felt it was important to understand how long we would expect these chronic multipliers to continue for and whether it is realistic that the relative impact of a fracture on HRQoL at 2 years will be the same as the impact at 10 years. The company reported evidence of long-term impact of fractures from several studies in response to clarification question B17D.[9] This included studies by Adachi et al. 2011, Blomfeldt et al. 2005 and Ekström et al. 2009.[127-129] These studies found that EQ-5D utilities remained lower than pre-fracture utilities after 3-, 5- and 2-years post-fracture, respectively.[127-129] Although the ERG could only see evidence up to 4 years in the Blomfeldt publication, it did show a continuing steady decline in utility between months 4, 12, 24, and 48 post-displaced femoral neck fracture, which could be likely to continue.[128] Ekström shows a steady-state lower post-fracture utility at months 4, 12 and 24 post- subtrochanteric fracture.[129] These studies suggest that a long-term effect of fracture on HRQoL could be appropriate. The same lifetime chronic multiplier assumption was made in TA464 and ID901, so could be considered an accepted approach. The ERG could not test the impact of this assumption as they could not change the VBA code in the model and the company declined to add an option for a reduced duration of chronic multipliers in the model.

4.2.8.2 Disutility values

Utility decrements were included for patients experiencing GIAEs whilst on oral bisphosphonate treatment. A fixed QALY decrement of 0.0075 was applied at the start of the treatment without adjustment for baseline health utility for 3% of patients when starting treatment with an oral bisphosphonate, in line with the assumptions included in Davis et al. 2015 as part of NICE TA464.[95]

ERG comment: It is unclear how this disutility was calculated in TA464 but given the size of the disutility and the percentage of patients it is applied to it is unlikely to have a large impact on results.

At clarification the company provided the option to include CV AEs in the model. A multiplier for QoL after a CV event was estimated based on a Swedish study by Lindgren et al. 2007,[130] which estimated a QoL loss of 0.075 (multiplier 0.910) during the first year after CV event. For the second and following years, the multiplier was assumed to be 0.95 due to lack of data.[9]

4.2.9 Resources and costs

The following cost categories were included in the analysis: drug acquisition costs, drug administration costs, disease management costs, costs associated with fractures (i.e., hip fractures, vertebral fractures, and NHNV fractures), long-term care costs after a hip fracture, and costs for the treatment of GIAEs.

4.2.9.1 Drug acquisition costs

The drug acquisition costs for romosozumab are £427.75 per set of two pre-filled disposable 1.17 ml injections of 90 mg/ml at list price or [redacted] including the Patient Access Scheme (PAS) discount, resulting in an annual cost of £5,133 at list price, or [redacted] including the PAS discount. The drug acquisition cost for alendronate at list price is £0.96 per pack with four tablets of 70 mg, or £13 annually. The cost of the comparators used in the scenario analyses are provided in Table 4.20 below.

Table 4.20: Drug acquisition costs

Drug	Annual drug cost	Pack size and cost	Method of administration	Dosing interval	Source
Treatments used in base-case analysis					
Romosozumab[a]	List: £5,133 PAS: £[redacted]	Injection, 90 mg/ml, consisting of two pre-filled disposable injections List: £427.75 PAS: £[redacted]	SC	QM	BNF 2021,[131] PAS
Alendronate	£13	70mg 4-tablet pack (£0.96)	Oral	QW	BNF 2021[131]
Treatments used in scenario analyses					
Teriparatide[b] (Forsteo)	£3,547	Injection, 250 micrograms/ ml, net price 2.4 ml prefilled pen=£271.88	SC	1 day	NHS indicative price 2021
Teriparatide[b] (Movymia)	£3,065	Injection, 250 micrograms/ ml, net price 2.4 ml prefilled pen (£235)	SC	QD	NHS indicative price 2021
Denosumab	£371	One pre-filled disposable injection (£180)	SC	Q6M[c]	BNF 2021[131]
Risedronate	£68	35mg 4-tablet pack (£18.88)	Oral	QW	BNF 2021[131]
Zoledronate	£85	Generic zoledronate 5[c] mg/ 100ml infusion bag	IV	Yearly	BNF 2021[131]
Raloxifene	£50	28-tablet pack (£3.81)	Oral	QD	BNF 2021[131]

Based on Table 31 in the CS.[1]

[a] Romosozumab is a 12 month course of treatment; [b] Treatment with teriparatide is limited to 24 months during a lifetime.[132]; [c] The ERG corrected the information from the CS, as explained in the ERG comment in Section 4.2.4 of the ERG report.

BNF = British National Formulary; CS = company submission; IV = intravenous; NHS = National Health Service; PAS = Patient Access Scheme; QD = once daily; QM = once monthly; Q6M = once every 6 months; QW = once weekly; SC = subcutaneous

4.2.9.2 Drug administration costs

No drug administration costs were included for romosozumab, which the company justified by referring to their plans to set up a Patient Support Programme (PSP) that includes homecare service, an adherence support program, and training of injection techniques. Administration costs are not included for alendronate since it is administered orally.

Drug administration costs were included in the model only for patients receiving denosumab or zoledronate. For patients receiving denosumab these consist of two nurse visits per year, which were valued at £9.50 assuming a 15 minute visit and using a unit cost of £38 per hour as provided by the Personal Social Services Research Unit (PSSRU) 2020.[133] For patients receiving zoledronate the administration cost was valued at £160 assuming the same cost as for delivery of chemotherapy and using the NHS National Tariff Workbook 2020/2021 (HRG code SB12Z; Deliver Simple Parental Chemotherapy at First Attendance).[134]

ERG comment: During the clarification phase, the ERG requested the inclusion of administration costs for romosozumab (i.e., representing a situation where the PSP is not in place) and all relevant comparators. The company responded by providing the results of scenario analyses that included the following administration costs in addition to those included in the original analyses: 12 nurse visits per year for romosozumab and 365 nurse visits per year for teriparatide. Nurse visits were valued at £9.50 (i.e., the same as above). For their base-case analysis, the ERG assumed a situation where the PSP has not (yet) been implemented and includes the costs for administration (i.e., 12 nurse visits) of romosozumab. The ERG performed a scenario analysis where it is assumed that the PSP is in place, and in which the costs of administration are applied in isolation as well as in combination with the assumption that persistence with romosozumab is 90%. The latter scenario was included since it is likely that the PSP leads to improvements in persistence with romosozumab.

4.2.9.3 Disease management costs

Disease management costs that were included in the model consist of BMD measurements and physician (GP) visits. BMD measurements were modelled at a frequency of once per two years and were valued at £40 using the NHS National Tariff Workbook 2020/2021 (RD50Z, DXA scan).[134] Physician visits for the monitoring of osteoporosis therapies were modelled at a frequency of once per year and were valued at £39 using the unit cost for a 9.22 minutes consultation as provided by the PSSRU 2020.[133]

ERG comment: The inclusion of costs for BMD measurements and physician visits was in line with Borgström et al. 2006 and Jönssen et al. 2011.[103, 135] However, other economic evaluations have included the costs of physician visits at a frequency of twice per year instead of only once, as indicated in the ESCEO/IOF recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019 and as used in Hiligsmann et al. 2020. The ERG preferred base-case analysis therefore assumed a frequency of twice per year for physician visits.

4.2.9.4 Fracture costs

The costs of hip, vertebral, and NHNV fractures during the first year after a fracture were sourced from a study by Gutiérrez et al.,[136, 137] and updated to 2020 using the consumer price indices (CPIs) as provided by the Office for National Statistics (ONS).[138] This resulted in cost estimates of £13,203, £2,897, and £2,131 for the first year after a hip, vertebral, or NHNV fracture, respectively. The costs of fractures in subsequent years were sourced from Davis et al. 2016,[95] and updated to 2020 using the CPIs as provided by the ONS.[138] These were only applied to hip and vertebral fractures at £115 and £361,

respectively. The costs of long-term care were included as recommended by the ESCEO/IOF recommendations for the conduct of economic evaluations in osteoporosis by Hiligsmann et al. 2019 and in line with TA464.[11, 78] In line with TA464, the probabilities of discharge to institutional care by age group were sourced from Najayan et al. 2014.[11, 117] The cost of long-term care in a nursing home was sourced from Hernlund et al. 2013,[80] and updated to 2020 using the CPIs as provided by the ONS,[138] which resulted in a daily cost of £112.

ERG comment: The first-year costs of hip, vertebral and NHNV fractures that were sourced from Gutiérrez et al., were based on the total costs.[136, 137] However, Gutiérrez et al. also provide the incremental costs of patients with fractures relative to matched controls. Since the incremental costs are more specific for the costs that are associated with the fracture and the model does not include additional costs of patients who do not sustain fractures beyond the disease management costs, the ERG considers it more appropriate to use the incremental costs for their base-case analysis. A similar approach based on incremental costs was also used in TA464 and ID901.[11, 87] The incremental first year costs provided by Gutiérrez et al., updated to 2019/2020 using the NHSCII as provided by the PSSRU 2020,[133] are £5,369 for a hip fracture, £1,465 for a vertebral fracture, and £877 for a NHNV fracture. A disadvantage of using these incremental cost estimates is that these do not include rehabilitation costs, which were included in the total cost for hip fracture used in the company's analyses.

The ERG notes that in TA464 a unit cost for long-term care was used and that was based on the assumptions that 1) equal proportions of patients who are discharged to long-term care go to nursing homes and residential care homes, 2) costs in the private sector are applicable (i.e., since the private sector provides 78% of places), and 3) that 36% of care is self-funded.[11] Using the unit costs as provided in PSSRU 2020,[133] £836 per week for private sector nursing homes and £620 per week for private sector residential care, the daily cost of long-term care can be estimated as 0.64 x (620+836) / 2 / 7 = £67. The ERG preferred to use this value for their base-case analysis.

4.2.9.5 Adverse event cost

Adverse event costs were applied to GIAEs at £40, based on a combination of the unit cost for a physician visit (see above) and a course of proton pump inhibitors (generic ranitidine, 300 mg tablets) at £0.90, sourced from the British National Formulary (BNF) January 2021.[131]

The company included the option to include CV AEs for those patients without a contraindicating history. The company identified the direct costs of CV events from a SLR from 2018.[139] This study estimated hospitalisation costs, outpatient referrals, primary care visits and medications of MI, stroke, unstable angina, heart failure, transient ischemic attack, and coronary artery bypass graft/percutaneous transluminal coronary angioplasty (CABG/PTCA), using hospital episodes statistics (HES) and CPRD data.[140] The estimated mean costs in the first 6 months after the first CV event was £4,594.16 in 2014 prices (£4993.85 in 2020, inflated using the indexes in Table 63 of the response to request for clarification[9]). Mean annualised cost in month 7 to 36 was £2,262.92 in 2014 prices (inflated to £2,459.79 in 2020 prices). The economic model was built to accommodate first and subsequent year costs, respectively. Therefore, the month 1 to 6 costs were applied in the first year and the month 7 to 36 costs were applied annually in every subsequent year until end of model time horizon or death. The company noted that this is likely to be a conservative approach as the first-year cost may be slightly overestimated in the model, since the majority costs likely occur closely to the event.[141]

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

Table 5.1 shows the deterministic CE results of the company's base-case analysis. All results are discounted and based on the confidential PAS price for romosozumab. Given that there are two relevant comparators, results are reported in a full incremental way. Pairwise ICERs of ROMO/ALN vs. each of the comparators (ALN and no treatment) are also reported for completeness. Results indicated that no treatment is dominated by ALN. Compared to ALN, ROMO/ALN accrued [redacted] incremental QALYs at [redacted] additional costs. Therefore, the ICER was £16,660 per QALY gained.

Table 5.1: Company base-case deterministic cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				3,747
ALN	[redacted]	10.014	[redacted]	[redacted]	0.021	[redacted]		16,660
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.031	[redacted]	16,660	

Based on Table 38 of the CS.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN

ALN = alendronate; CS = company submission; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = patient access scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

The disaggregated discounted costs are shown in Table 5.2 for the comparison vs. ALN and in Table 5.3 for the comparison vs. no treatment.

Table 5.2: Disaggregated cost results (ROMO/ALN vs. ALN)

Cost item	Cost intervention (ROMO/ALN)	Cost comparator (ALN)	Increment	Absolute increment	Absolute increment (%)
Hospitalisation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nursing home	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 1st treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 2nd treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 165 of CS Appendix J.[8]

ALN = alendronate; CS = company submission; ROMO/ALN = romosozumab-to-alendronate

Table 5.3: Disaggregated cost results (ROMO/ALN vs. no treatment)

Cost item	Cost intervention (ROMO/ALN)	Cost comparator (no treatment)	Increment	Absolute increment	Absolute increment (%)
Hospitalisation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Outpatient	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nursing home	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 1st treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Drug cost: 2nd treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment management	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 166 of CS Appendix J.[8]
ALN = alendronate; CS = company submission; ROMO/ALN = romosozumab-to-alendronate

The company did not present disaggregated results for QALYs but reported differences in fracture events over 10 years between treatment arms, which is the main driver of the difference in QALYs produced by the model. These results are displayed in Tables 5.4 and 5.5.

Table 5.4: Summary of number of fracture events over 10 years, ROMO/ALN vs. ALN

Fracture type	Fracture events intervention (ROMO/ALN)	Fracture events comparator (ALN)	Difference
Hip	[redacted]	[redacted]	[redacted]
Vertebral	[redacted]	[redacted]	[redacted]
NHNV	[redacted]	[redacted]	[redacted]
Any	[redacted]	[redacted]	[redacted]

Based on Table 163 of CS Appendix J.[8]
ALN = alendronate; CS = company submission; NHNV = non-hip, non-vertebral; ROMO/ALN = romosozumab-to-alendronate

Table 5.5: Summary of number of fracture events over 10 years, ROMO/ALN vs. no treatment

Fracture type	Fracture events intervention (ROMO/ALN)	Fracture events comparator (no treatment)	Difference
Hip	[redacted]	[redacted]	[redacted]
Vertebral	[redacted]	[redacted]	[redacted]
NHNV	[redacted]	[redacted]	[redacted]
Any	[redacted]	[redacted]	[redacted]

Source: Table 164 in CS Appendix J.[8]
ALN = alendronate; CS = company submission; NHNV = non-hip, non-vertebral; ROMO/ALN = romosozumab-to-alendronate.

Overall, the technology is modelled to affect QALYs by:

- Reducing the incidence of fractures.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments, and
- Reducing costs associated to a decreased number of fractures.

5.2 Company's sensitivity and scenario analyses

5.2.1 Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) in which all input parameters were sampled simultaneously from their corresponding probability distributions over 1,000 iterations. The input parameters and the probability distributions used in the PSA can be seen in Table 36 of the CS.[1] The main distributional assumptions for the model parameters highlighted by the company are described below:

- Drug unit costs are assumed to be fixed and, therefore, they are not sampled in the model. For all the other cost parameters, a lognormal distribution with a standard error of 25% of the base-case value was assumed.
- Utility multipliers for hip, vertebral and NHNV fractures were sampled from a lognormal distribution with standard errors based on study data.
- Persistence on treatment and proportions of patients going to long-term care after a hip fracture were sampled from a beta distribution.
- Risk ratios for treatment efficacy were sampled from a normal distribution. Standard errors were based on the trial data and/or NMA.

The average PSA results are summarised in Table 5.6, and presented on a CE plane in Figure 5.1, from which a CE acceptability curve (CEAC) was calculated and plot in Figure 5.2. Both the CE-plane and CEAC plots are based on the pairwise comparisons vs. ROMO/ALN.

Table 5.6: Company base-case probabilistic cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER** (£/QALY)
No treatment	[redacted]	NR	[redacted]	Dominated by ALN				3,976*
ALN	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]		14,537
ROMO/ALN	[redacted]	NR	[redacted]	[redacted]	NR	[redacted]	14,537	

Based on Table 39 of the CS.[1]

* Not the same as in the CS, probably due to rounding of QALYs; ** All pairwise ICERs are calculated vs. ROMO/ALN

ALN = alendronate; CS = company submission; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; NR = not reported; PAS = patient access scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

The average PSA results are in line with the deterministic ones shown in Table 5.1. Also, in the PSA no treatment is dominated by ALN, and the ICER for the comparison ROMO/ALN vs. ALN was £14,537 per QALY gained. The lower PSA ICER is the result of both lower incremental costs and higher incremental QALYs for ROMO/ALN vs. ALN. As shown in Figure 5.1, at the threshold of

£30,000 per QALY gained, the estimated probability that ROMO/ALN is a cost-effective alternative to ALN was [redacted] and [redacted] compared to no treatment.

Figure 5.1: Probabilistic sensitivity analysis cost effectiveness plane (PAS price for romosozumab)

Based on Figure 16 of the CS.[1]
Note: mind the axes of the CE-plane; they are not presented in their most common form (x-axis for incremental QALYs and y-axis for incremental costs)
ALE = alendronate; CE = cost effectiveness; CS = company submission; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROM = romosozumab

Figure 5.2: Probabilistic sensitivity analysis cost effectiveness acceptability curve (PAS price for romosozumab)

Based on Figure 17 of the CS.[1]
ALE = alendronate; CS = company submission; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROM = romosozumab; WTP = willingness to pay

5.2.2 Deterministic sensitivity analysis

The company also conducted deterministic sensitivity analyses (DSAs) where key parameters were individually varied at lower and upper bounds of values that were deemed plausible by the company. These are summarised in Table 5.7. Note that parameters like the starting age in the model, the length of the time horizon or the duration of the offset time are usually not included in the DSA but in scenario analyses.

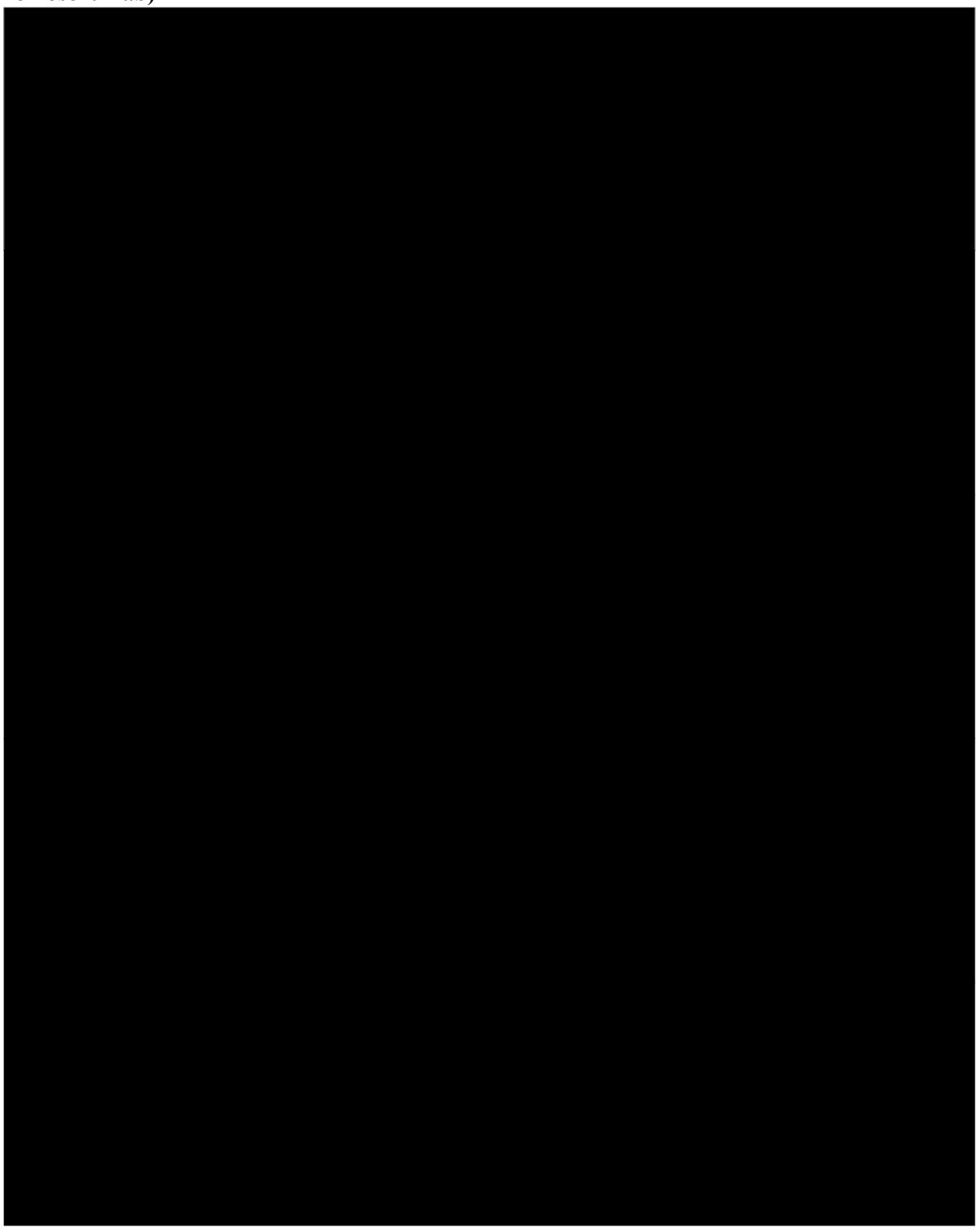
Table 5.7: Parameters and values included in the company's DSA

Parameter	Values	ERG comment
Start age	50, 60, 70 and 80 years	Scenario analyses (not DSA)
Model time frame	5, 10, 15 and 20 years	Scenario analyses (not DSA)
Fixed offset time	1, 2, 3, 4, 5 and 6 years	Scenario analyses (not DSA)
Utility multiplier for hip, vertebral and NHNV fracture in the first year following fracture	95% CI	Agree, evidence based
Utility multiplier for hip, vertebral and NHNV fracture in the second and following years after fracture	95% CI	Agree, evidence based
Direct medical cost first year after fracture	±25% of base-case	Agree, commonly used
Direct medical cost second and following years after hip and vertebral fracture	±25% of base-case	Agree, commonly used
Daily cost for long term care after hip fracture	±25% of base-case	Agree, commonly used
RRs for hip, vertebral and NHNV fractures for romosozumab	95% CI	Agree, evidence based
RRs for hip, vertebral and NHNV fractures for alendronate	95% CI	Agree, evidence based
Persistence multiplier for romosozumab	±25% of base-case	Arbitrary
Persistence multiplier for alendronate	±25% of base-case	Arbitrary

Based on Table 40 in CS.[1]
CI = confidence interval; CS = company submission; DSA = deterministic sensitivity analysis; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; NHNV = non-hip, non-vertebral; RR = relative risk

The results of the DSAs are presented in Table 41 in the CS.[1] This table shows pairwise ICERs for the comparisons ROMO/ALN vs. ALN and ROMO/ALN vs. no treatment for all scenarios defined in Table 5.7. For the comparison vs. no treatment, all ICERs were below the £20,000 per QALY gained threshold (or ROMO/ALN was dominant), except for the following scenarios: start age 50 years (ICER was £28,721), start age 60 years (ICER was £31,642) and time horizon 5 years (ICER was £49,862). The ICER was more sensitive to changes for the comparison vs. ALN. The results for this comparison were summarised by the company in the form of a tornado diagram as shown in Figure 5.3. This shows that the model results are sensitive to varying the time horizon, persistence, start age, changes in treatment effect of romosozumab/alendronate and alendronate alone, and utility multipliers for hip, vertebral and NHNV fracture.

Figure 5.3: DSA tornado diagram for romosozumab/alendronate vs. alendronate (PAS price for romosozumab)



Based on Figure 18 of the CS.[1]
CS = company submission; DSA = deterministic sensitivity analysis; GI = gastrointestinal; ICER = incremental cost effectiveness ratio; NHNV = non-hip, non-vertebral; QALY = quality-adjusted life year; RR = relative risk; vert = vertebral

5.2.3 Scenario analysis

The company conducted several scenario analyses in which the CE of ROMO/ALN was analysed against comparators that were not included in the base-case analysis. A summary of these scenarios is provided in Table 5.8. Scenario analyses 1 to 9 were based on the NMA using the ITT populations of ARCH and FRAME. Scenario 10 was based on the NMA using the EU label-matched populations from ARCH and FRAME. A patient population with a recent MOF, 74 years at treatment start, T-score of −2.9 and fracture risk corresponding to approximately 30% based on FRAX was assumed for scenarios 1 to 10. Scenario 11 was conducted for the comparison of ROMO/ALN vs. denosumab, as in scenario 7, but assuming a patient population at a higher risk of fracture. In particular, the assumed patient population for this scenario consisted of 74-year-old women, with a recent MOF and a T-score of −[redacted] and an approximately 10-year probability of MOF of [redacted]% according to FRAX. The results of the scenario analysis are presented in Table 5.9. All results include PAS price for ROMO. Results showed that ROMO/ALN was dominant or ICERs below £20,000 per QALY gained except for the comparisons against denosumab in the base-case population (£35,400 in scenario 7) and in the higher risk population (£27,509 in scenario 11).

Table 5.8: Summary of company scenario analyses

Scenario	Comparison	Treatment length	Offset	NMA efficacy source
1	ROMO/ALN vs. ALN	ROMO: 12m ALN: 48m vs. ALN: 60m	Dynamic	ITT population
2	ROMO/ALN vs. TERI	ROMO: 12m ALN: 48m vs. TERI: 24m	Dynamic	ITT population
3	ROMO/ALN vs. TERI	ROMO: 12m ALN: 48m vs. TERI: 18m	Dynamic	ITT population
4	ROMO/ALN vs. TERI biosimilar/ALN	ROMO: 12m ALN: 48m vs. TERI bio: 18m ALN: 42m	Dynamic	ITT population
5	ROMO/ALN vs. TERI/ALN	ROMO: 12m ALN: 48m vs. TERI: 18m ALN: 42m	Dynamic	ITT population
6	ROMO/ALN vs. RAL	ROMO: 12m ALN: 48m vs. RAL: 60m	Dynamic	ITT population
7	ROMO/ALN vs. DENO	ROMO: 12m ALN: 48m vs. DENO: 60m	ROMO: Dynamic DENO: 12m	ITT population
8	ROMO/ALN vs. RIS	ROMO: 12m ALN: 48m vs. RIS: 60m	Dynamic	ITT population
9	ROMO/ALN vs. ZOLE	ROMO: 12m ALN: 48m vs. ZOLE: 60m	Dynamic	ITT population
10	ROMO/ALN vs. ALN	ROMO: 12m ALN: 48m vs. ALN: 60m	Dynamic	ARCH EU*
11	ROMO/ ALN vs. DENO**	ROMO: 12m ALN: 48m vs. DENO: 60m	ROMO: Dynamic DENO: 12m	ITT population

Source: Table 42 and 43 in CS.[1]

* ARCH-EU label-matched population used in NMA. ** Scenario conducted for a population with a higher risk of fracture.

Note: For DENO, the company assumed a clinical effect limited to within 6 months after stopping treatment.[27, 111] The company explained that chronic treatment with DENO is necessary when used as the subsequent treatment after ROMO for this combination to provide optimal benefits to patients; or alternatively a further treatment switch to a bisphosphonate after the DENO treatment period would be required. Therefore, a 1-year fixed offset time was applied to DENO.

ALN = alendronate; CS = company submission; DENO = denosumab; EU = European Union; ITT = intention-to-treat; m = months; NMA = network meta-analysis; RAL = raloxifene; RIS = risedronate, ROMO = romosozumab; TERI = teriparatide, ZOLE = zoledronate

Table 5.9: Company scenario analyses results (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER (£/QALY)*
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	9.993	[redacted]	Dominated by DENO				3,747
RALO (6)	[redacted]	9.998	[redacted]	Dominated by DENO				Dominated
RIS (8)	[redacted]	10.013	[redacted]	Dominated by DENO				12,518
ALN (1)	[redacted]	10.014	[redacted]	Dominated by DENO				16,660
TERI (3)	[redacted]	10.021	[redacted]	Dominated by DENO				Dominated
TERI (2)	[redacted]	10.023	[redacted]	Dominated by DENO				Dominated
TERI/ALN (5)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
TERI biosimilar/ALN (4)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
ZOLE (9)	[redacted]	10.026	[redacted]	Dominated by DENO				17,176
DENO (7)	[redacted]	10.034	[redacted]					35,400
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.011	[redacted]	35,400	
Scenario 10								
ALN	[redacted]	10.013	[redacted]					
ROMO/ALN	[redacted]	10.043	[redacted]	[redacted]	0.030	[redacted]	17,690	
Scenario 11								
DENO	[redacted]	9.800	[redacted]					
ROMO/ALN	[redacted]	9.813	[redacted]	[redacted]	0.013	[redacted]	27,509	

Based on Tables 44, 45 and 46 of the CS.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; CS = company submission; DENO = denosumab; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = patient access scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

5.2.4 Additional scenario analysis requested by the ERG

Some areas of uncertainty were identified by the ERG during the clarification phase, resulting in the company conducting additional scenario analyses requested by the ERG in the clarification letter.[9] The uncertainties explored by the company in these additional scenarios are the following:

- **Imminent risk of fracture**. The ERG requested a scenario where the imminent risk of fracture was removed from the analysis. This was because the ERG considers it unclear whether the reduction in fracture risk from treatment, estimated from the ARCH ITT population, corresponds to a population with imminent risk of fracture (see Key issue 1). The company indicated that this scenario should not be considered relevant for this appraisal because it does not accurately represent the romosozumab target patient population. While this might be the case, given the uncertainties previously mentioned, the ERG considers that this scenario provides relevant information. Results are shown in Table 5.10. The ICER increased by £18,523 compared to the base-case ICER of £16,660 per QALY gained.
- **Incidence of vertebral fractures**. Following the discussion in Section 4.2.6.1, the ERG asked the company to conduct a scenario analysis where the results from Singer et al. 1998 for vertebral fractures were assumed.[81] The company explained that the vertebral fracture incidences estimate from this study are generally not considered to be reliable. For that reason, the results from this scenario should not be considered relevant for this appraisal because, according to the company, it likely underestimates the risk of clinical vertebral fractures and therefore underestimates the CE of romosozumab. However, given the uncertainties concerning the company's approach described in Section 4.2.6.1, the ERG considers that this scenario has informational value, in only for providing an upper limit for the ICER with regards to the uncertainty about the incidence of vertebral fractures. Results are shown in Table 5.11. The ICER increased by £14,052 compared to the base-case ICER of £16,660 per QALY gained.
- **Treatment effect estimated from an alternative NMA**. The company considered that results for alendronate vs. placebo were similar in both the NICE and the CS NMA. The ERG argued that this is a subjective statement seeing the values presented in Table 4.8, especially for the values shown for teriparatide. This raised concerns about the validity/credibility of the NMA results. Hence, the ERG asked the company to provide results based on the NICE NMA. The company concluded that CE scenarios based the NICE NMA are not appropriate for this appraisal because the underlying evidence base for such NMA was outside the licensed indication for romosozumab. However, given the uncertainties concerning the company's NMA, as highlighted in Key issue 3, the ERG considers this a valid scenario. The results in Table 5.12, show that the ICER was similar to the ICER in the company base-case.
- **Persistence**. Persistence assumptions were identified as one of the most important drivers of the CE results. Concerns regarding the company base-case assumptions on persistence and how these could bias the results in favour of romosozumab were explained in Section 4.2.6.2. Based on these, the ERG asked the company to explore three additional scenarios in which 1) persistence was assumed to be as in the ARCH trial for romosozumab and the alendronate, 2) persistence on romosozumab was assumed to be equal to persistence on teriparatide and 3) an unrealistic scenario with 100% compliance in both intervention and comparator. Again, the company indicated that these scenarios are not relevant for this appraisal. In particular, for the first scenario, the company emphasised that persistence inputs derived from clinical trials are known to differ substantially from real-world persistence of osteoporosis patients and are at high risk to misrepresent the CE of romosozumab. The ERG agrees with this and as explained in Section 4.2.6.2, considers that by using trial-based persistence for romosozumab vs. real life

persistence for alendronate, there is indeed a high risk that the CE of romosozumab is misrepresented in the company base-case. Even though it is known that real-life persistence will be lower than in trial settings, at least this scenario would provide a fair comparison. For the second scenario, the company considered that persistence to romosozumab is unlikely to be equal to teriparatide's persistence given the difference in administration frequency (romosozumab is given monthly and teriparatide is given daily). While the ERG acknowledged that this might be the case, the company has not provided evidence to support this assumption. Hence, the relevance of this scenario. Finally, even if it seems clear that a scenario based on 100% persistence is unrealistic, the results of this scenario can still be relevant for decision-making. Results are shown in Tables 5.13 to 5.15. In all scenarios the ICER increased compared to the base-case, especially in the first one where the ICER was almost £40,000 higher.

- **Alternative risk of death for hip and vertebral fractures**. The company run a scenario where the relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al. 2007 (UK setting).[142] The relative risks in the second and following years for hip and vertebral fractures, and first year for NHNV fractures, were assumed to be the same as in the base-case. Results are shown in Table 5.16. This had a minor impact on the CE results.
- **CV adverse events**. The ERG asked the company to include in the analysis CV AEs according to the incidence in the ARCH trial and relevant disutilities and costs. The company indicated that the results of this scenario can be considered conservative for romosozumab since the CV occurrence rates for romosozumab and alendronate were chosen from the study where the imbalance between these two treatments was greatest (ARCH) and subsequent year costs are applied every year after the CV event until the end of the modelled time horizon or death. The decision not to select or pool any other romosozumab studies (FRAME, STRUCTURE, McClung) where the CV event rate for romosozumab was lower than in ARCH to derive CE results of this scenario means that the results should be considered to be extremely conservative, and for illustrative purposes only. Nonetheless, the ERG considers that since the efficacy results are based on ARCH it is appropriate that AE evidence is based on ARCH. Results are shown in Table 5.17. The ICER increased by £2,840 compared to the base-case.
- **Drug administration costs**. The company ran a scenario including drug administration costs (i.e., for subcutaneous injections) when the PSP is not in place for romosozumab, as well as for the relevant comparators that are used in scenario analyses. The cost (£9.5 per administration) was based on a 15-minute visit (based on £38 per hour for GP nurse contact time). PSSRU Unit Costs of Health and Social Care 2020 10.2 Nurse (GP practice). Unit costs available 2019/2020 based on 1,573 hours per year, which includes 225 working days minus sickness absence (8 days) and any training/study days as reported for all NHS staff groups. In the scenario analysis, romosozumab is associated with 12 SC injections days (i.e., 24 injections) per year administered by a nurse; teriparatide 365 injections per year and denosumab two injections per year. Results are shown in Table 5.18. All ICERs increased (moderately) compared to those shown in Table 5.9.

Table 5.10: Company scenario with fracture recency removed (no imminent risk) cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	10.044	[redacted]	Dominated by ALN				12,688

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ALN	[redacted]	10.055	[redacted]	[redacted]	0.011	[redacted]		35,183
ROMO/ALN	[redacted]	10.074	[redacted]	[redacted]	0.019	[redacted]	35,183	

Based on Table 44 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.11: Company scenario with vertebral fracture incidences from Singer et al. 1998 cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	10.069	[redacted]	Dominated by ALN				8,967
ALN	[redacted]	10.075	[redacted]	[redacted]	0.006	[redacted]		30,712
ROMO/ALN	[redacted]	10.087	[redacted]	[redacted]	0.012	[redacted]	30,712	

Source: Based on Table 45 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate.

Table 5.12: Company scenario using efficacy of ALN vs. placebo from NICE NMA cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				4,219
ALN	[redacted]	10.013	[redacted]	[redacted]	0.020	[redacted]		17,069
ROMO/ALN	[redacted]	10.045	[redacted]	[redacted]	0.032	[redacted]	17,069	

Based on Table 47 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; NMA = network meta-analysis; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.13: Company scenario with persistence data based on ARCH for all treatments cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				646

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ALN	[redacted]	10.034	[redacted]	[redacted]	0.041	[redacted]		54,340
ROMO/ALN	[redacted]	10.051	[redacted]	[redacted]	0.017	[redacted]	54,340	

Based on Table 53 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.14: Company scenario with romosozumab persistence equal to teriparatide persistence cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN				10,016
ALN	[redacted]	10.014	[redacted]	[redacted]	0.021	[redacted]		38,295
ROMO/ALN	[redacted]	10.032	[redacted]	[redacted]	0.018	[redacted]	38,295	

Based on Table 54 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate.

Table 5.15: Company scenario with 100% persistence for all treatments cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.993	[redacted]	Dominated by ALN and ROMO/ALN				Dominated
ALN	[redacted]	10.045	[redacted]	[redacted]	0.052	[redacted]		20,989
ROMO/ALN	[redacted]	10.072	[redacted]	[redacted]	0.027	[redacted]	20,989	

Based on Table 55 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.16: Company scenario using relative risk of death for hip and vertebral fractures during the first year were based on the study by van Staa et al. 2007 cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.981	[redacted]	Dominated by ALN				3,824
ALN	[redacted]	10.000	[redacted]	[redacted]	0.019	[redacted]		16,728

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
ROMO/ALN	[redacted]	10.031	[redacted]	[redacted]	0.031	[redacted]	16,728	

Based on Table 59 of the clarification letter response.[9]
Note: ICERs are not exactly the same as those reported by the company, possibly due to rounding.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.17: Company scenario including cardiovascular adverse events cost effectiveness results (PAS price for romosozumab)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
No treatment	[redacted]	9.966	[redacted]	Dominated by ALN				5,075
ALN	[redacted]	9.986	[redacted]	[redacted]	0.020	[redacted]		19,500
ROMO/ALN	[redacted]	10.013	[redacted]	[redacted]	0.027	[redacted]	19,500	

Based on Table 60 of the clarification letter response.[9]
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; ROMO/ALN = romosozumab-to-alendronate

Table 5.18: Company scenario analyses results including cost for subcutaneous administrations (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	9.993	[redacted]	Dominated by DENO				5,123
RAL (6)	[redacted]	9.998	[redacted]	Dominated by DENO				Dominated
RIS (8)	[redacted]	10.013	[redacted]	Dominated by DENO				14,953
ALN (1)	[redacted]	10.014	[redacted]	Dominated by DENO				19,434
TERI (3)	[redacted]	10.021	[redacted]	Dominated by DENO				Dominated
TERI (2)	[redacted]	10.023	[redacted]	Dominated by DENO				Dominated
TERI/ALN (5)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
TERI biosimilar/ALN (4)	[redacted]	10.025	[redacted]	Dominated by DENO				Dominated
ZOLE (9)	[redacted]	10.026	[redacted]	Dominated by DENO				21,129
DENO (7)	[redacted]	10.034	[redacted]					43,000
ROMO/ALN	[redacted]	10.044	[redacted]	[redacted]	0.010	[redacted]	43,000	

Source: Based on Table 61 of the clarification letter response.[9]
Note: It is unclear why Table 61 of the clarification letter response provides different QALYs/LYG than those in Table 5.10 since only costs are supposed to change.
* All pairwise ICERs are calculated vs. ROMO/ALN
ALN = alendronate; DENO = denosumab; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RAL = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

5.2.5 Conclusions from company's sensitivity and scenario analyses

The modelling assumptions that have the greatest effect on the ICER are:

- Treatment persistence.
- Start age of the population.
- Model time horizon.
- Treatment effect of romosozumab followed by alendronate and alendronate alone.
- Utility multipliers for hip, vertebral and NHNV fracture.
- Comparator choice (denosumab).

5.3 Model validation and face validity check

Validation efforts conducted on the economic model were discussed in the validation section of the CS (B.3.10).[1] In general, the model adheres to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] A comparison between the recommended key modelling aspects and the assumption made on the romosozumab model is provided in Table 47 of the CS.[1] Modelling assumptions were also validated by leading UK experts at an advisory board held by the company in 2017.

Most of the validation efforts discussed in the CS referred to those conducted by NICE's PRIMA (Preliminary Independent Model Advice service) in 2017.[122, 143] PRIMA assessed the appropriateness of the conceptual model, model verification (through black-box testing), reproducibility and made suggestions on how to improve the model's transparency and usability. The complete PRIMA report was presented as part of the CS. Furthermore, the company explained that the model has also undergone thorough quality control by Quantify Research, including performing multiple verification and validation tests, as well cross-validating the results with another in-house osteoporosis model.

The company also mentioned that the same model has been used in two published peer-reviewed manuscripts,[73, 74] and in the reimbursement submissions of romosozumab in Sweden (TLV) and

Scotland (SMC).[104, 125] Although not explicitly stated, it is assumed that the model might have also passed quality controls previous to publication and/or during the reimbursement assessments.

Additionally, the number of fractures predicted by the CE model was validated using a Swedish cohort study of women 50 years and older with fracture identified in the National Patient Register. Details of the Swedish registry study can be found elsewhere.[86] Since BMD data were available from three large hospitals in Sweden, a comparison between the model and real-world fracture incidences adjusted for risk factors such as age and BMD was possible. For this comparison, the romosozumab model was populated with Swedish population incidences and used the Swedish version of FRAX. Using the registry data, the incidence of fracture (all types) was predicted for 5-year follow-up with a multiple-failure model. The 10-year incidence was calculated using the non-parametric single-failure model. These were compared with the incidence predicted by the health economic model. The results of this comparison can be seen in Table 5.19. The CE model predicted approximately [redacted]% higher 5-year incidence than the incidence estimated from the registry data. The company considered that this can be is mainly explained by the fact that vertebral fractures are at risk of being underreported in register data. Ten-year incidence was calculated using register data for women 55 to 90 years with MOF and unknown BMD. However, the same population cannot be completely reproduced in the CE model, which makes this comparison of limited value. In the CE model, the fracture risk is likely to be higher than the fracture risk for the average Swedish population 55 to 90 years with unknown BMD. This is shown in Table 5.20. However, the extent to which the 10-year risk predicted by the model are comparable to the risk observed in real life is unknown.

Table 5.19: Validation of simulated fracture risks using Swedish register data

Source	Outcome	Women with MOF**, age 74, unknown BMD	Women with MOF**, age 74, T-score -2.9	Women with MOF, age 55-90**, unknown BMD
Register study	5-year cumulative incidence of new fracture (disregarding type)*	34.6% (1a)	52.5% (1b)	
CE model***	5-year cumulative incidence of new fracture	[redacted]% (1a)	[redacted]% (1b)	
Register study	10-year non-parametric cumulative incidence of a new fracture (single failure model)			37.6%
CE model***	10-year risk of a new fracture (single failure model)	[redacted]%	[redacted]%	

Based on Table 48 of the CS.[1]

* Predicted incidence based on a multiple failure model; ** At baseline; *** Excess mortality of fracture set to 100%. The CE model adjusts mortality for comorbidities, i.e., mortality unrelated to the fracture. This adjustment cannot be made in the register data; therefore, excess mortality was set to 100% in the model for better comparison.

BMD = bone mineral density, CE = cost effectiveness, CS = company submission; MOF = major osteoporotic fracture

Finally, in response to clarification question B27,[9] the company provided a comparison of the distribution of fractures in the Swedish real-world study vs. the distribution of fractures in the CE model. In the Swedish real-world study, out of the 231,769 patients with at least one fracture, 7,656 patients (3.3%) had a third fracture over approximately 5.5 years of maximum follow-up data.[85] The CE model estimated 4.4% of patients had a third fracture over 5 years. The company explained that these values are not strictly comparable since in the Swedish data, the first fracture could have happened at some point during the 5.5 years of follow-up, meaning that not all patients would have enough follow-up time to have developed a second or a third fracture.

ERG comment: The model adheres in general to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] Since 2017, the model has been involved in several iterations of quality assessment including the NICE PRIMA. In line with this assessment, the ERG considers that review would be facilitated if calculations were performed in the model work sheets, instead of being hard coded in VBA. As explained in Section 4.2.2, the VBA code was initially password protected because the FRAX algorithm is confidential. After clarification, the company provided the rest of the VBA which was reviewed by the ERG. The VBA code was well structured and sufficient comments were provided to understand the flow of the code. In reviewing the model and the VBA code, the ERG noted the following issues:

- In the 'State trace' sheet of the model the proportions of patients with a first NHNV fracture (i.e. column M) always remains zero, whereas from the second cycle onwards there is a non-zero proportion of patients with a second NHNV fracture. The ERG could not trace the source of this issue.
- After running the model with the 'Trackers summary' enabled the ERG noted that the means of outpatient costs do not match with the means of outpatient cost on the 'Results' sheet. From scrutinizing the VBA code in module mRunModel.bas, it appears that t_IterCost (comparator, 3) is not updated (lines 3264-3270) for costs in year 2 and more after hip and vertebral fracture. If this is indeed the cause, it seems that it does not impact the overall results.
- Also, in the 'Trackers summary' the drug costs and treatment management costs always remain zero but not in the 'Results' sheet. The ERG could not trace the cause of this. Note that the means of other costs, LYs and QALYs did match between the 'Trackers summary' and the 'Results' sheet.
- In the module mRunModel.bas an error was found in line 2065. In the formula PrevFx = PrevFx + t_fx(comparator, 1) + t_fx(comparator, 1) + t_fx(comparator, 3) the second 't_fx(comparator, 1)' should read 't_fx(comparator, 2)'. It is not clear to the ERG to what extent this impacts the results.

An additional point the ERG would like to emphasise is the model running time. Despite the added complexity of microsimulation compared to standard cohort models, the model seems to be extremely demanding regarding the computational power needed to run within reasonable time. Even a deterministic run would take more than 20 minutes. This makes the validation process extra difficult and for this reason, the ERG was not able to validate the results of some of the scenarios presented by the company. In particular, the ERG did not succeed in running any PSA. Sometimes the model would stop running after a few PSA iterations and most of the times Excel would crash. The default settings of 500,000 iterations for the inner loop and 1,000 for the outer loop projected a running time of more than 2 weeks to finish, which in practice can be deemed as unfeasible. Given this practical issue, the ERG would like to suggest the company to conduct an analysis to estimate the minimal PSA loop sizes that would provide reliable results in a minimum running time and to re-consider the programming of the model in order to make it computationally more efficient.

As explained in detail in the ERG comment in Section 4.2.6 (baseline fracture incidence), there is uncertainty regarding the validity of the incidences of hip, vertebral and NHNV fractures, relating to the aspects:

- The company used a study that dates from 1998 by Singer et al.[81] as the main source of input values.
- The company referred to a study by van der Velde et al. 2016 to confirm the stability of hip fracture incidence over time but which had substantially lower incidence rates than Singer et al. 1998.[81, 82]
- The company referred to a study by Kanis et al. 2001 to confirm the similarity between ratios of vertebral to hip fractures in Sweden and the UK.[90] The ERG could not confirm that a comparison between ratios of vertebral to hip fractures in Sweden and the UK was included in Kanis et al. 2001.
- For the different types of fractures that were included in the estimates of the incidence of NHNV fractures that were sourced from Singer et al. 1998, the company referred to van der Velde et al. 2016 to confirm the stability over time and similarity of findings from both studies.[81, 82] However, the ERG could not confirm the stability over time and the similarity of findings for all types of fractures that were included.

Validation was presented against Swedish data only. The company indicated that it was not possible to perform the validation based on UK data, since detailed data on fractures and risk factors such as BMD were not available. Therefore, it is uncertain whether the validity results can be generalised to the UK.

Comparisons with other TAs were not presented. Therefore, it is not possible to quantify whether the results in the CS are in line with those in previous appraisals.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

6.1.1 Explanation of the company adjustments after the request for clarification

During the clarification phase, the ERG requested the company to correct errors on the 'PSA input' sheet that resulted in cells displaying '#N/A' and '#NUM!'. The company provided a corrected version of the model alongside their response to the ERG's clarification questions.

6.1.2 Explanation of the ERG adjustments

The changes that the ERG can make (to the model received with the response to the clarification letter) can be subdivided into the following three categories (according to Kaltenthaler et al. 2016[144]):

- Fixing errors (correcting the model where the company's electronic model is unequivocally wrong).
- Fixing violations (correcting the model where the ERG considers that the NICE reference case, scope or best practice has not been adhered to).
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred).

In the current assessment, only matters of judgement played a role. After the proposed changes were implemented in the company's model, additional scenario analyses were explored by the ERG in order to assess the impact of alternative assumptions on the CE results.

6.1.2.1 Fixing errors

No errors were corrected by the ERG in the model provided in response to the clarification letter. Note that the ERG was granted access to a version of the model in which the VBA code was unprotected to facilitate validation by the ERG. However, the company was unable to perform exhaustive quality assurance on the "unprotected" version of the model and asked the ERG to use the model received with the response to the clarification letter to conduct all ERG scenarios. As a consequence, the ERG was not able to change any of the model VBA code, regardless of whether this was with the purpose of fixing errors or testing alternative assumptions.

6.1.2.2 Fixing violations

No violations were applicable to this appraisal.

6.1.2.3 Matters of judgement

The ERG's preferences regarding alternative assumptions led to the following changes to the company base-case analysis:

- Romosozumab persistence (i.e., at 6 and 12 months) was changed from 90% to 80% (see Section 4.2.6.2).
- Alendronate persistence was changed as follows: for alendronate after romosozumab the ERG used estimates for persistence with oral bisphosphonates in non-naïve patients from Morley et al. 2020 and for alendronate alone the ERG used estimates for persistence with oral bisphosphonates in naïve patients from Morley et al. 2020 (see Section 4.2.6.2).[100]
- Only excess mortality for hip fractures (and not for other types of fractures) was included in the analysis (see Section 4.2.6.4).

- Daily costs of long-term care were changed from £112 to £67 (see Section 4.2.9).
- The ERG changed the input parameter values for the costs associated with fractures from £13,203 to £5,369 for hip fractures, from £2,897 to £1,465 for vertebral fractures, and from £2,131 to £877 (see Section 4.2.9).
- Cardiovascular events which occurred in patients who did not have a history of MI or stroke were included in the analysis (see Section 4.2.7).
- Costs for administration of romosozumab (and for the comparators denosumab and teriparatide) that are applicable as long as the PSP is not in place were included in the analysis (see Section 4.2.9).
- The frequency of physician visits was changed from once per year to twice per year (see Section 4.2.9).
- General population mortality input parameter values were updated to the most recent UK National Life Tables (see Section 4.2.6.4).

The overview of the changes and the bookmarks for the justification of the ERG changes are presented in Table 6.1.

Table 6.1: Company and ERG base-case preferred assumptions

Base-case preferred assumptions		Company	ERG	Justification for change
Persistence with romosozumab		90%	80%	Section 4.2.6.2
Persistence with alendronate	**Alendronate after romosozumab**	85% of persistence with denosumab as reported in Karlsson et al. 2015[97]	Morley et al. 2020 persistence with oral BPs in non-naïve patients[100]	Section 4.2.6.2
	Alendronate alone	Li et al. 2012[96]	Morley et al. 2020 persistence with oral BPs in naïve patients[100]	
Excess mortality following fractures		Included for hip, vertebral and NHNV fractures	Included for hip fractures only	Section 4.2.6.4
Daily costs of long-term care		£112	£67	Section 4.2.9
Costs associated with fractures	**Hip**	£13,203	£5,369	Section 4.2.9
	Vertebral	£2,897	£1,465	
	NHNV	£2,131	£877	
Cardiovascular events		Not included	Included	Section 4.2.7
Romosozumab administration costs (PSP)		Not included (PSP in place)	Included (PSP not in place)	Section 4.2.9
Frequency of physician visits		Once per year	Twice per year	Section 4.2.9
General population mortality		2012-2014 UK National Life Tables	2017-2019 UK National Life Tables	Section 4.2.6.4

BP = bisphosphonates; ERG = Evidence Review Group; NHNV = non-hip, non-vertebral; PSP = Patient Support Programme; UK = United Kingdom

6.1.3 Additional scenarios conducted by the ERG

The ERG conducted a series of scenario analyses to explore the impact of key assumptions and uncertainties within the CE analyses. These uncertainties were related to the inclusion of comparators other than alendronate alone, removal of the imminent risk, persistence, the PSP, costs associated with fractures, utilities, AEs, treatment effect waning, and excess mortality associated to fractures.

6.1.3.1 Scenario set 1: other comparators

The ERG performed scenario analyses using the same comparators defined by the company in Section 5.2.3: teriparatide, zoledronate, denosumab, risedronate, raloxifene and no treatment.

6.1.3.2 Scenario set 2: imminent risk removed

To address the uncertainty regarding the relevant population for this appraisal, as discussed for example in Section 4.2.3, the ERG performed a set of scenario analyses where the "imminent risk" of fracture was removed from the analysis. This set of scenarios was performed with all comparators as in scenario set 1.

6.1.3.3 Scenario set 3: persistence

To address the uncertainty regarding assumptions on persistence with osteoporosis therapies, the ERG performed the following set of scenario analyses:

- No distinction is made between alendronate naïve (i.e., patients receiving alendronate alone) and non-naïve patients (i.e., patients receiving alendronate after romosozumab). Thus, this scenario assumes the same persistence for patients receiving alendronate after romosozumab and alendronate alone, both persistence estimates based on persistence with oral BPs in Morley et al. 2020 in the 'All patients' (i.e., naïve patients and non-naïve patients pooled) population.[100]
- An analysis where it is assumed that persistence with romosozumab is the same as in the company base-case; i.e., 90% instead of 80%.
- A scenario was also conducted assuming persistence for romosozumab as per the ERG base-case and persistence for alendronate as per the company base-case.
- The persistence scenarios requested at clarification were also repeated on the ERG base-case, including using the ARCH trial persistence for both romosozumab and alendronate; assuming the persistence on romosozumab was equal to that of teriparatide and assuming 100% persistence for all treatments.

6.1.3.4 Scenario set 4: patient support programme in place

To address the uncertainty regarding the impact on CE results following the implementation of the company's plans to set up the PSP, the ERG performed a set of scenario analyses where no administration costs are assumed for romosozumab and where the assumption of no administration costs is combined with the assumption of 90% persistence with romosozumab.

6.1.3.5 Scenario set 5: costs associated with fractures

To address the uncertainty regarding the costs associated with fractures, the ERG performed a scenario analysis assuming total health care costs associated with fractures from Gutiérrez et al. 2011 and 2012 (i.e., the same as in the company base-case analysis, which also includes rehabilitation costs for hip fractures), instead of the incremental costs of patients with fractures vs. those without (as in the ERG base-case analysis, which does not include rehabilitation costs) from the same sources.[136, 137]

6.1.3.6 Scenario set 6: utility multipliers

Although the application of utility multipliers for fracture events has been a common approach in previous osteoporosis appraisals[11, 87], the multipliers differ somewhat across appraisals. Therefore, scenarios using the alternative sets of multipliers (shown in Table 4.19 of this report) were conducted to examine the impact on results.

6.1.3.7 Scenario set 7: adverse events

The ERG included those CV AEs which occurred in patients without a history of MI or stroke in their base-case as an imbalance was observed in the ARCH trial. A scenario was also conducted where these CV AEs were excluded.

6.1.3.8 Scenario set 8: treatment effect waning

The ERG run a scenario in which 4 years of full treatment effect was assumed followed by a waning in effect for one more year. The fracture risk ratios assumed for the fifth year were the following: [redacted] for hip fracture, [redacted] for vertebral fracture and [redacted] for NHNV fractures. The dynamic offset was equal to 5 years.

6.1.3.9 Scenario set 9: excess mortality associated to fractures

Following ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] the ERG base-case included excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNV fractures were also explored by the ERG.

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.2.1 Results of the ERG preferred base-case scenario

The ERG preferred base-case incremental CE results, provided in Table 6.2, indicate that the total costs associated with romosozumab (12 months) followed by alendronate (48 months) were estimated at [redacted] and the total costs associated with alendronate alone (60 months) were estimated at [redacted], indicating an incremental cost of [redacted]. Total QALYs associated with romosozumab (12 months) followed by alendronate (48 months) were estimated at [redacted] and total QALYs associated with alendronate alone (60 months) were estimated at [redacted], indicating an incremental number of [redacted] QALYs gained. These results indicate an estimated ICER of £483,750 per QALY gained.

It should be highlighted that in the ERG base-case, the incremental LYGs are negative. This is due to the inclusion of serious CV AEs in the ERG base-case, which occurred more frequently in the romosozumab arm than in the alendronate alone arm, and which had an impact on mortality.

Table 6.2: ERG preferred base-case deterministic cost effectiveness results (discounted, PAS)

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Romosozumab followed by alendronate	[redacted]	10.048	[redacted]	[redacted]	-0.002	[redacted]	483,750
Alendronate alone	[redacted]	10.050	[redacted]				

Based on the ERG preferred version of the electronic model.[1]
Note: The results of the comparison vs. no treatment are reported in Section 6.2.2.1 of the ERG report.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year.

As mentioned in Section 5.3, the ERG was unable to run a PSA for its preferred base-case analysis. However, given the deterministic ICER and assuming that the PSA ICER would be in line with this one, the probability that romosozumab is considered cost effective at a threshold of £30,000 compared to alendronate is likely to be [redacted]%.

6.2.2 Results of the ERG additional exploratory scenario analyses

6.2.2.1 Scenario set 1 results: other comparators

The results of scenario analyses set 1, using various alternative comparators, are provided in Table 6.3. These indicate that the relevant comparison is zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab followed by alendronate, show that all ICERs are above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 1 month, teriparatide 24 months and teriparatide followed by alendronate, which are dominated by romosozumab followed by alendronate; and the comparison against zoledronate, which is dominant.

Some counterintuitive results were observed when teriparatide was involved as a comparator treatment. It is unclear why the sequence teriparatide (or biosimilar) would result in less QALYs than teriparatide alone (even if teriparatide alone is given for 24 months and for 18 months as part of the sequence). If this would be the case, it would seem irrational to treat patients with the sequence when teriparatide alone is more beneficial. Also, note that this was not observed in the results presented by the company in Table 5.9. Therefore, the ERG explored this potential issue a bit further and run an "extreme" scenario in which teriparatide 18 months was compared with teriparatide 18 months followed by alendronate, but with persistence on alendronate equal to zero. In this scenario, teriparatide alone resulted in [redacted] QALYs and the sequence with alendronate at zero persistence resulted in [redacted] QALYs. Thus, the sequential treatment provided more QALYs even when persistence on the second treatment on the sequence was equal to zero. A similar scenario was run but with romosozumab instead of teriparatide and the same effect on QALYs was observed. The ERG was not able to find the source for these inconsistencies, which might need further confirmation from the company. It is also unclear why the sequence with teriparatide biosimilar would result in more QALYs than the sequence with commercial teriparatide. This is likely due to both sequences being informed by different NMAs.

Table 6.3: Scenario set 1 results: other comparators (PAS price for romosozumab)

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
No treatment	[redacted]	10.0440	[redacted]	Dominated by RIS				£44,288
RALO (6)	[redacted]	10.0397	[redacted]	Dominated by RIS				£37,000
RIS (8)	[redacted]	10.0493	[redacted]	Dominated by ALN				£226,438
TERI (3) 18m	[redacted]	10.0509	[redacted]	Dominated by ALN				Dominated by ROMO/ALN
ALN (1)	[redacted]	10.0500	[redacted]					£483,750
DENO (7)	[redacted]	10.0532	[redacted]	Dominated by ZOLE				£1,088,000
TERI/ALN (5)	[redacted]	10.0516	[redacted]	Dominated by TERI bio/ALN				Dominated by ROMO/ALN
TERI biosimilar/ALN (4)	[redacted]	10.0516	[redacted]	Extendedly dominated by ROMO/ALN				£228,000
TERI (2) 24m	[redacted]	10.0515	[redacted]	Dominated by ROMO/ALN				Dominated by ROMO/ALN
ROMO/ALN	[redacted]	10.0484	[redacted]	Dominated by ZOLE				ZOLE dominates
ZOLE (9)	[redacted]	10.0492	[redacted]	[redacted]	-0.001	[redacted]	£47,583	

Based on the ERG preferred version of the electronic model.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; DENO = denosumab; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

6.2.2.2 Scenario set 2 results: imminent risk removed

The results of scenario analyses set 2, with the imminent risk removed, are provided in Table 6.4. These indicate that the relevant comparison is zoledronate vs. alendronate, with an ICER of £121,730 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab followed by alendronate, show that all ICERs are well above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 18 months, and teriparatide followed by alendronate, which are dominated by romosozumab followed by alendronate; and the comparisons against zoledronate and denosumab, which are dominant. The same counterintuitive results discussed in the previous section were also observed in this set of scenarios. Furthermore, it also seems counterintuitive that raloxifene was dominated by no treatment. However, this can be explained by looking at fracture risk ratios presented in Table 4.11. Therefore, the model results for this scenario seem consistent with the NMA input but the ERG is concerned about the validity of the value provided by the NMA.

Table 6.4: Scenario set 2 results: imminent risk removed

Technologies (scenario number)	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)	Pairwise ICER* (£/QALY)
Scenarios 1 – 9 (including no treatment)								
RALO (6)	[redacted]	10.0508	[redacted]	Dominated by no treatment				£76,548
No treatment	[redacted]	10.0543	[redacted]	Dominated by RIS				£98,965
RIS (8)	[redacted]	10.0591	[redacted]	Dominated by ALN				£667,218
TERI (3) 18m	[redacted]	10.0595	[redacted]	Dominated by TERI/ALN				Dominated by ROMO/ALN
TERI/ALN (5)	[redacted]	10.0601	[redacted]	Dominated by TERI bio/ALN				Dominated by ROMO/ALN
TERI bio/ALN (4)	[redacted]	10.0601	[redacted]	Dominated by ALN				£3,454,305
ROMO/ALN	[redacted]	10.0581	[redacted]	Dominated by ALN				
ALN (1)	[redacted]	10.0599	[redacted]					ALN dominates
TERI (2) 24m	[redacted]	10.0609	[redacted]	Dominated by DENO				£11,872,642
DENO (7)	[redacted]	10.0619	[redacted]	Dominated by ZOLE				DENO dominates
ZOLE (9)	[redacted]	10.0596	[redacted]	[redacted]	-0.0003	[redacted]	£121,730	ZOLE dominates

Based on the ERG preferred version of the electronic model.[1]

* All pairwise ICERs are calculated vs. ROMO/ALN.

ALN = alendronate; DENO = denosumab; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALY = quality-adjusted life year; RALO = raloxifene; RIS = risedronate; ROMO/ALN = romosozumab-to-alendronate; TERI = teriparatide; ZOLE = zoledronate

6.2.2.3 Scenario set 3 results: persistence

The results of scenario analyses set 3, using various alternative assumptions and inputs for persistence, are provided in Table 6.5. These scenario results demonstrate the substantial and varied impact of different persistence assumptions on results. Using the persistence estimates from Morley et al. 2020,[100] based on all patients for persistence with oral BPs, substantially increased the incremental QALYs and reduced the ICER by approximately £400,000 per QALY gained. Assuming 90% persistence for romosozumab resulted in an ICER approximately mid-way between the ERG base-case and the company base-case at £267,533 per QALY gained. The scenario assuming romosozumab persistence per the ERG base-case and comparator persistence per the company base-case and the scenario assuming all treatments had persistence of 100% resulted in similar substantial increases in incremental QALYs and ICERs of approximately £40,000 per QALY gained (a decrease of approximately £443,000 in the ICER. Scenarios assuming persistence data based on trial data for all treatments and assuming romosozumab persistence equal to that of teriparatide resulted in negative incremental QALYs for romosozumab followed by alendronate, resulting in the treatment being dominated by alendronate.

Table 6.5: Scenario set 3 results: persistence

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Morley 2020 'All patients' for persistence with oral BPs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	81,333
90% persistence with romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	267,533
Romo persistence per ERG BC; Comparators per company BC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,315
Persistence based on trial data for all treatments	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Romo dominated
Romo persistence equal to teriparatide persistence	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Romo dominated

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
All treatments 100% persistence	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	40,539

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.4 Scenario set 4 results: patient support programme in place

The results of scenario analyses where it is assumed that the PSP is in place, are provided in Table 6.6. Assuming no administration costs for romosozumab had a minor impact on the results. In the scenario where the same assumption was combined with 90% persistence with romosozumab the ICER was almost halved.

Table 6.6: Scenario set 4 results: patient support programme in place

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
No admin. costs for romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	471,250
No admin. costs + 90% persistence with romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	260,533

Based on the ERG preferred version of the electronic model.[1]
admin. = administration; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.5 Scenario set 5 results: costs associated with fractures

The results of scenario analyses set 5 are provided in Table 6.7. In this scenario the total health care costs associated with fractures from Gutiérrez et al. 2011 and 2012 are applied, instead of the incremental costs of patients with fractures vs. those without from the same sources.[136, 137] The impact of this assumption on the model results was minimal.

Table 6.7: Scenario set 5 results: costs associated with fractures

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
ERG base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Scenario 5: total health care costs associated with fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	482,750

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.6 Scenario set 6 results: utility multipliers

The results of the utility multiplier scenarios are provided in Table 6.8. The ICER was very sensitive to the multipliers applied as the incremental QALYs in the ERG base-case are so small and, therefore, changes to incremental QALYs have a large impact on the ICER. Using the TA646 multipliers approximately doubled the incremental QALY gain to [redacted] from [redacted], which led to a substantial reduction in the ICER to £258,000 from £483,750. Conversely, using the multiplier from ID901led to a small decrease of approximately [redacted] in the incremental QALYs, but still increased the ICER by approximately £70,000 per QALY gained.

Table 6.8: Scenario set 6 results: utility multipliers

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Base-case multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
TA464 multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	258,000
ID901 multipliers	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	552,857

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.7 Scenario set 7 results: CV AEs

The results of the scenario in which CV AEs were removed from the ERG base-case are shown in Table 6.9. Removing the CV AEs led to a decrease in incremental costs and an increase in incremental QALYs, resulting in a decrease of approximately £173,000 in the ICER.

Table 6.9: Scenario set 7 results: CV AEs

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
CV AEs included (ERG)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
No CV AEs (company)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	310,917

Based on the ERG preferred version of the electronic model.[1]
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.8 Scenario set 8 results: treatment effect waning

Results for the treatment effect waning scenario are displayed in Table 6.10. The scenario in which 4 years of full treatment effect was assumed followed by a waning in effect for one more year resulted in a slight increase in incremental costs, and a slight reduction in incremental QALYs, which led to a substantial increase in the ICER of approximately £70,000 per QALY gained.

Table 6.10: Scenario set 8 results: treatment effect waning

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
No effect waning (BC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
4 years full effect then 1 year waning	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	554,714

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; QALY(s) = quality-adjusted life year(s).

6.2.2.9 Scenario set 9 results: excess mortality associated to fractures

Results for excess mortality scenarios are displayed in Table 6.11. The ERG base-case assumed excess mortality after hip fracture only. Including excess mortality also after vertebral fracture decreased the ICER by approximately £130,000 per QALY gained, due to an increase in incremental QALYs. The further addition of excess mortality due to NHNV had almost no impact on the ICER.

Table 6.11: Scenario set 9 results: excess mortality associated to fractures

Scenario	Romosozumab followed by alendronate		Alendronate alone		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Hip only (ERG BC)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750
Hip and vertebral	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	355,273
Hip, vertebral and NHNV	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	354,545

Based on the ERG preferred version of the electronic model.[1]
BC = base-case; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; Incr. = incremental; NHNV = non-hip non-vertebral; QALY(s) = quality-adjusted life year(s)

6.3 ERG preferred assumptions

Tables 6.12 and 6.13 show the step-by-step changes made by the ERG to the company base-case alongside the cumulative and one-by-one impact of each change on the results, respectively. The change with the largest impact (by far) on the results was sourcing alendronate persistence estimates from Morley et al. 2020.[100] This highlights the importance of persistence parameters on the CE results. Other changes like including CV events in the analysis had a large impact on the cumulative base-case, because the ICER now was very sensitive given the small incremental QALYs, but not when this change is applied alone. The following three changes, when applied in isolation, resulted in an ICER that increased from below to above £20,000 per QALY gained: assuming 80% persistence with romosozumab (i.e., instead of 90%), assuming a daily cost of long-term care of £67 (i.e., instead of £112), and assuming incremental costs associated with fractures (i.e., of patients with fractures vs. those without, instead of total health care costs). The other changes, when applied in isolation, also resulted in increased ICERs but still remained below £20,000 per QALY gained.

Table 6.12: Incremental impact of ERG preferred assumptions (cumulative)

Preferred assumption (Section in ERG report)	Romosozumab 12 months / alendronate 48 months		Alendronate 48 months		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	262,429
+ Excess mortality only for hip fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,000
+ Daily LTC costs £67	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,000
+ Incremental fracture costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	303,750
+ CV events included	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	473,375
+ No PSP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	485,875
+ 2 GP visits per year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	484,250
+ UK general population mortality 2017 - 2019	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	483,750

Based on the ERG preferred version of the electronic model.[1]
CV = cardiovascular; ERG = Evidence Review Group; GP = general practitioner; ICER = incremental cost effectiveness ratio; Incr. = incremental; LTC = long-term care; PSP = patient support programme; QALY(s) = quality-adjusted life year(s)

Table 6.13: Incremental impact of ERG preferred assumptions (one-by-one)

Preferred assumption (Section in ERG report)	Romosozumab 12 months / alendronate 48 months		Alendronate 48 months		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,660
+ 80% for persistence romosozumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	21,483
+ Morley et al. 2020 for persistence alendronate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	162,391
+ Excess mortality only for hip fractures	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,185
+ Daily LTC costs £67	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,476
+ Incremental fracture costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20,398
+ CV events included	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,500
+ No PSP	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,680
+ 2 GP visits per year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,117
+ UK general population mortality 2017 - 2019	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,903

Based on the ERG preferred version of the electronic model.[1]
CV = cardiovascular; ERG = Evidence Review Group; GP = general practitioner; ICER = incremental cost effectiveness ratio; Incr. = incremental; LTC = long-term care; PSP = patient support programme; QALY(s) = quality-adjusted life year(s).

6.4 Conclusions of the cost effectiveness section

The selection of databases searched was very comprehensive. Full details of the database searches, including the database name, host platform and date searched, were clearly and transparently reported. Overall, the ERG does not have any major concerns regarding the searches but notes that no searches were conducted to identify health-state utility values (see Section 4.1.1 for more details), it is unclear whether empirical studies estimating utility values in this condition were missed as only included economic evaluations were searched for utility values.

The company developed a "de novo" Markov microsimulation model in Microsoft Excel. The model structure appears appropriate in general. However, the ERG's ability to step through and evaluate the model functionality was hindered by the fact that all model calculations are done in background VBA code that could not be changed. Therefore, the ERG was unable to assess the functionality of the model or to make changes to assumptions beyond simple input parameters. The CE analysis was performed in line with the NICE Reference case in terms of perspective, time horizon and discounting.[77]

The population in the Final Scope by NICE is defined as "Postmenopausal women with severe osteoporosis at high risk of fracture", in line with romosozumab marketing authorisation. The modelled population in the CS is assumed to consist of patients who are at imminent risk of another fragility fracture i.e. have had a MOF within the prior 24 months. An important difference between the ARCH ITT population and the modelled population is that ARCH included patients who previously sustained a fracture regardless of recency, whereas for the modelled population it is assumed that a previous fracture was sustained within 24 months prior to the start of treatment. In the ARCH ITT population, [redacted] of patients suffered a MOF within 24 months prior to randomisation. The differences between the definition in the NICE final scope, the ITT population from ARCH that was used to inform treatment effectiveness inputs for the company's base-case analysis, and the definition of the modelled population in the CS, present a key issue of uncertainty. It is not clear whether the term 'high risk' as used in the definitions in the NICE final scope and EMA marketing authorisation corresponds to the same definition that is used in the literature for the categorisation of fracture risk to guide choice of treatment. It is, therefore, uncertain whether the ITT population results are representative for the population in the CS and whether these are generalisable to the target population of romosozumab.

The modelled intervention consisted of a 12-month course of romosozumab, followed by a 48-month course of alendronate. The comparators that were used in the company base-case consist of a 60-month course of alendronate and no treatment. Additionally, the company included additional comparators (teriparatide, denosumab, risedronate, zoledronate, and raloxifene) as scenario analyses. All treatments considered by the company were listed in the NICE scope, except for ibandronic acid, for which the company identified no trials at its licensed dose and, therefore, it could not be included in the analyses. Given the uncertainty regarding the relevant population for this appraisal in Section 2.1, as previously described, and the lack of clarity of current guidelines, there is also uncertainty regarding the appropriateness and relevance of the included comparators. In particular, if high risk is differentiated from very high risk, then alendronate might be the most appropriate comparator, but if high risk includes very high risk, then other comparators might be appropriate.

In the model, the risk of fractures in patients with severe osteoporosis who had a MOF in the prior 24 months is estimated using three components: general population risk of fractures, increased risk of fractures associated with osteoporosis, imminent risk of subsequent fractures following an index fracture. The general population risk of hip fractures was sourced from Singer et al. 1998 and the same source was used to estimate the incidence of vertebral fractures using the ratio of hip to vertebral fractures from a Swedish study.[81, 83] To estimate the incidence of NHNV fractures, Singer et al. 1998 was used for forearm fractures and the same approach that was used to estimate the incidence of vertebral fractures was applied to the other types of fractures that are included in NHNV fractures.[81] No changes were applied by the ERG, but the ERG did note some uncertainty regarding the validity of estimates of fracture incidence that was related to the stability over time of fracture incidence and the assumption that ratios between different types of fractures as found in Sweden also apply to the UK. The increased risk of fracture due to osteoporosis, relative to the general population, was estimated using FRAX whilst excluding prior fracture as a clinical risk factor. Finally, the additional risk of experiencing a subsequent fracture after an index fracture was based on the maximum of the 'imminent

risk', sourced from Söreskog et al. 2020,[85] or the additional risk from FRAX whilst including prior fracture as a risk factor.

Efficacy estimates for romosozumab vs. alendronate were based on ARCH data, by reconstructing patient-level data from published Kaplan-Meier curves and then fitting parametric distributions in order to calculate time-dependent hazard rates. These (survival data) analyses were not presented by the company. While the methods used seem appropriate, the ERG cannot assess whether the distributions were properly fitted and cannot explore the impact of using alternative distributions on the model results. In analyses vs. other comparators, efficacy was estimated using an NMA in which treatment effects were estimated on the trial ITT population. Limitations of the NMA were discussed in the clinical effectiveness sections of the report (e.g., Section 3.6).

The company modelled persistence with osteoporosis therapies based on the assumption that real-world persistence with romosozumab would equal persistence as found in ARCH and that persistence with alendronate following romosozumab would be 85% of real-world persistence with denosumab from Li et al. 2012. Persistence with alendronate alone was also based on Li et al. 2012.[96] The ERG identified Morley et al. 2020 [100] as a more recent source of persistence estimates that is effectively an update of the study by Li et al. 2012 (both based on GPRD), that they preferred to use for their base-case analysis to inform persistence with alendronate, using estimates from non-naïve patients for alendronate after romosozumab and estimates from naïve patients for alendronate alone. This change, when applied in isolation of the other ERG changes, resulted in the largest impact on the CE results and increased the ICER by nearly ten-fold.

The company assumed that anti-fracture efficacy persists for a period of time (offset time) after treatment is discontinued in patients with osteoporosis.[102] A dynamic offset time equal to time on treatment is assumed for the base-case. During the offset time the fracture risk reduction is assumed to decline linearly to zero. The efficacy of the last treatment given to the patient in the sequence was used for the offset time. This approach is recommended by the ESCEO and IOF guidelines, and has been used in other published health economic studies and romosozumab HTA submissions to the SMC and TLV.[78, 103, 104] Therefore, the ERG considers the company's approach appropriate. Scenarios with fixed offset time can be deemed as exploratory. As described in Key issue 2, scenarios with shorter duration of the dynamic offset of the treatment effect could be of interest of being explored. However, the ERG was unable to run this type of scenarios, which are expected to increase the ICER. Finally, the ERG would like to note that residual effects for zoledronate could be longer than those assumed by the company.[112] The ERG was unable to change the model to incorporate this assumption. Cost effectiveness results including zoledronate as comparator might be underestimating the ICER (even though in these scenarios zoledronate was dominant over romosozumab).

Mortality is captured in the model in three ways: age-specific mortality of the general population (all-cause mortality), relative risk capturing excess mortality of the disease and co-morbidity adjustment factor.[1] It is unclear why the company used UK Life Tables from 2012 to 2014. In the ERG base-case, the most recent version (2017 to 2019) was used.[115] When a patient sustains a fracture, the relative risk of death compared with the non-fractured population is applied to the normal population risk, and the relative risk was down-adjusted to 30% to adjust for higher frailty (i.e., increased risk of death due to other reasons than the fracture itself) in the fractured population.[1, 94, 114] The company included in the base-case mortality related to hip, clinical vertebral, and NHNV fractures. Following the expert reviewers comments to the ESCEO/IOF recommendations for economic evaluations in osteoporosis,[78] the ERG prefers to include excess mortality after hip fractures only. Scenarios assuming excess mortality after vertebral fractures, and after NHNV fractures were explored by the ERG.

The company only included GI AEs associated with bisphosphonates in their base-case. An imbalance in serious adjudicated CV AEs was observed in the ARCH trial, which led to romosozumab being contraindicated for patients with previous MI or stroke. The company chose to exclude CV AEs from their base-case due to this contraindication. However, the ERG considered that those CV events which occurred in patients without a history of MI or stroke should be included as they would not be avoided by the contraindication. These CV events in patients without history were therefore included in the ERG base-case.

Utilities for fracture health states within the model were estimated using fracture multipliers from the international ICUROS study, multiplied with UK age adjusted general population utility values. Separate multipliers were provided for hip, vertebral and NHNV fractures during the first (acute) and subsequent (chronic) years after fracture. This utility approach follows previous appraisals in osteoporosis, although some differences in multipliers were observed across appraisals. Multipliers from other available NICE appraisals were used in scenarios to examine the impact of differences on results. The ERG was unsure about the appropriateness of several assumptions in the utility analysis. In TA464, only one acute multiplier could be applied at any one time, while in this model two acute multipliers could be applied at once. Additionally, the ERG was unclear whether the assumption that chronic fracture multipliers were used for the remainder patients' lifetimes was supported by evidence. The company presented some evidence up to 5 years post fracture, but none beyond. However, the ERG was unable to test the impact of these assumptions, given that they could not access the VBA code in the validated version of the model on which analyses had to be conducted.

The following cost categories were included in the analysis: drug acquisition costs, drug administration costs, disease management costs, costs associated with fractures (i.e., hip fractures, vertebral fractures, and NHNV fractures), long-term care costs after a hip fracture, and costs for the treatment of GIAEs. The drug acquisition costs for romosozumab are £427.75 per set of two pre-filled disposable 1.17 ml injections of 90 mg/ml at list price or [redacted] including a PAS discount, resulting in an annual cost of £5,133 at list price, or [redacted] including the PAS discount. No drug administration costs were included for romosozumab, which the company justified by referring to their plans to set up a PSP that includes homecare service, an adherence support program, and training of injection techniques. Drug administration costs were included in the model only for patients receiving denosumab or zoledronate. However, since the PSP is not yet in place, the ERG preferred to include the costs for administration of romosozumab. For disease management costs, the ERG preferred the assumption that monitoring of osteoporosis therapies requires physician visits once a year to twice a year, in line with Hilligsmann et al. 2019.[78] The ERG preferred to use estimates of costs associated with fractures that were based on incremental costs of patients with fractures vs. patients without, rather than total costs of patients with fractures that were used by the company, in line with NICE TA464 and ID901.[11, 87] Lastly, the ERG preferred to use a different estimate of long-term costs based on the estimate as used in TA464.[11]

The company's deterministic base-case results indicate that romosozumab followed by alendronate is more costly and more effective than alendronate alone, with incremental QALYs of [redacted] and incremental costs of [redacted], resulting in an ICER of £16,660 per QALY gained. In the fully incremental analysis, no treatment was dominated by alendronate alone. The company's PSA results were more or less in line with their deterministic results, with a probabilistic ICER of £14,537 per QALY gained. At a threshold of £30,000 per QALY gained, the estimated probability that romosozumab is a cost-effective alternative to alendronate alone or no treatment is [redacted] and [redacted] respectively. The company's DSA shows that model results are sensitive to varying the time horizon, persistence, start age, changes in treatment effect of romosozumab/alendronate and alendronate alone, and utility multipliers for hip,

vertebral and NHNV fracture. Company scenario analyses highlighted the sensitivity of results to persistence assumptions and the removal of imminent risk in the calculation of fracture incidence.

The ERG base-case differed from the company base-case in a number of elements including: the assumed persistence rates for romosozumab and alendronate; assumed excess mortality after vertebral and NHNV fractures; incremental fracture and daily LTC costs; inclusion of CV AEs and PSP costs; number of GP visits per year and the source of UK general population mortality rates. The ERG change that had by far the most impact on results when applied in isolation was assuming the persistence estimate for alendronate from Morley et al. 2020. Which increased the company base-case ICER ten-fold, from £16,660 to £162,391 per QALY gained. The next most influential parameters reducing the daily LTC cost, assuming 80% persistence on romosozumab and reducing the incremental fracture costs, all of which when applied individually took the ICER over £20,000 per QALY gained.

The ERG deterministic base-case resulted in higher incremental costs ([redacted] vs. [redacted]) and substantially lower incremental QALYs ([redacted] vs. [redacted]) which resulted in a high ICER of £483,750. A PSA on the ERG base-case could not be run as the model continued to crash. However, given the size of the ICER, it is likely that at the usual threshold range of £20,000 to £30,000 per QALY gained, the probability of romosozumab being cost effective would be [redacted]%. Scenario analyses run on the ERG preferred assumptions showed that model results were most sensitive to assumed rates of persistence, however, scenarios surrounding utility multipliers, treatment effect waning, excess mortality due to fractures and inclusion of CV AEs and PSP also had large impacts on the ICER, which was very sensitive to changes in the small incremental QALYs. When various alternative comparators were included in the analysis, romosozumab was dominated by zoledronate. In this situation, the only relevant comparison was zoledronate vs. alendronate, with an ICER of £47,583 per QALY gained. All the other treatment options are either dominated or extendedly dominated. Pairwise comparisons against romosozumab, showed that all ICERs were above the threshold of £30,000 per QALY, except for the comparisons against teriparatide 18 months, teriparatide 24 months and teriparatide followed by alendronate, which are dominated by romosozumab.

Regarding validation, the model adheres in general to the recommendations on modelling in osteoporosis by ESCEO and IOF.[78] Since 2017, the model has been involved in several iterations of quality assessment including the NICE PRIMA. In line with this assessment, the ERG considers that review would be better facilitated if calculations were performed in the model work sheets, instead of being hard coded in VBA. Some discrepancies between the model results and the trackers summary were found, which could not all be traced to their source in the VBA code. An error was found with regards to the presence of previous fractures, and it is not clear is this has any impact on the results. Additionally, the model seems to be extremely demanding regarding the computational power needed to run within reasonable time. Even a deterministic run would take more than 20 minutes. This makes the validation process extra difficult and for this reason, the ERG was not able to validate the results of some of the scenarios presented by the company and did not succeed in running any PSA. The main concerns of the ERG relate to the validity of the baseline fracture incidences as noted above. Also, validation was presented against Swedish data only because for example UK data on fractures and risk factors such as BMD were not available. Therefore, it is uncertain whether the validity results can be generalised to the UK. Finally, comparisons with other technology appraisals were not presented. Therefore, it is not possible to quantify whether the results in the CS are in line with those in previous appraisals.

The same issues identified in the clinical effectiveness section are carried over in the economic analyses. The model results are affected by the limitations of the NMAs, and they should be interpreted in a

similar way as the results of the NMAs: with caution. If additional data are identified to reduce bias in the NMAs, this would also reduce the uncertainty around the model results. However, it is uncertain what the effect on the CE estimates might be.

In conclusion, in contrast to the company's base-case that resulted in an ICER of £16,660 per QALY gained, the ERG preferred base-case results in an ICER of £483,750 per QALY. This difference is mainly caused by different assumptions regarding the persistence with alendronate. The high value of the ICER can further be explained by the higher incremental costs ([redacted] vs. [redacted]) and substantially lower incremental QALYs ([redacted] vs. [redacted]) of the ERG preferred base-case vs. the company's base-case, respectively.

7 REFERENCES

[1] UCB Pharma Ltd. *Romosozumab for treating postmenopausal severe osteoporosis [ID3936]. Document B. Company evidence submission for romosozumab. National Institute of Health and Care Excellence. Single technology appraisal (STA)*: UCB Pharma Ltd, 2021

[2] National Institute for Health and Care Excellence. *Romosozumab for treating severe osteoporosis. Health Technology Appraisal. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/gid-ta10828/documents/final-scope

[3] Saag KG, Petersen J, Brandi ML, Karaplis AC, Lorentzon M, Thomas T, et al. Romosozumab or alendronate for fracture prevention in women with osteoporosis. *N Engl J Med* 2017;377(15):1417-27.

[4] UCB Pharma Limited. EVENITY 105 mg solution for injection in pre-filled pen (romosozumab). Summary of product characteristics [Internet]. electronic Medicines Compendium (eMC), 2021 [accessed 1.7.21]. Available from: https://www.medicines.org.uk/emc/product/10956/smpc

[5] Amgen Inc. *STRUCTURE. Clinical study report: 20080289. An open-label, randomized, teriparatide-controlled study to evaluate the effect of treatment with romosozumab in postmenopausal women with osteoporosis previously treated with bisphosphonate therapy. UCB Data on File*, 2015

[6] Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761-7.

[7] Alswat KA. Gender disparities in osteoporosis. *J Clin Med Res* 2017;9(5):382-7.

[8] UCB Pharma Ltd. *Romosozumab for treating postmenopausal severe osteoporosis [ID3936]. Appendices. Company evidence submission for romosozumab. National Institute of Health and Care Excellence. Single technology appraisal (STA)*: UCB Pharma Ltd, 2021

[9] UCB Pharma Ltd. *Romosozumab for treating severe osteoporosis [ID3936]. Clarification questions. National Institute of Health and Care Excellence. Single technology appraisal (STA). Company response to clarification*: UCB Pharma Ltd, 2021

[10] Golder S, Peryer G, Loke YK. Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *J Clin Epidemiol* 2019;113:36-43.

[11] National Institute for Health and Care Excellence. *Bisphosphonates for treating osteoporosis. NICE technology appraisal guidance 464 [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta464

[12] Bai H, Jing D, Guo A, Yin S. Randomized controlled trial of zoledronic acid for treatment of osteoporosis in women. *J Int Med Res* 2013;41(3):697-704.

[13] Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *N Engl J Med* 2007;356(18):1809-22.

[14] Morii H, Ohashi Y, Taketani Y, Fukunaga M, Nakamura T, Itabashi A, et al. Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. *Osteoporos Int* 2003;14(10):793-800.

[15] Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med* 2009;361(8):756-65.

[16] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021) [Internet]*: Cochrane, 2021 [accessed 1.7.21] Available from: https://training.cochrane.org/handbook

[17] Cosman F, Crittenden DB, Adachi JD, Binkley N, Czerwinski E, Ferrari S, et al. Romosozumab treatment in postmenopausal women with osteoporosis. *N Engl J Med* 2016;375(16):1532-43.

[18] Langdahl BL, Libanati C, Crittenden DB, Bolognese MA, Brown JP, Daizadeh NS, et al. Romosozumab (sclerostin monoclonal antibody) versus teriparatide in postmenopausal women with osteoporosis transitioning from oral bisphosphonate therapy: a randomised, open-label, phase 3 trial. *Lancet* 2017;390(10102):1585-94.

[19] Amgen Inc. *BRIDGE. Clinical study report: 20110174. A multicenter, randomized, double-blind, placebo-controlled study to compare the efficacy and safety of romosozumab with placebo in men with osteoporosis. UCB Data on File*, 2016

[20] Amgen Inc. *ARCH. Clinical study report: 20110142. A multicenter, international, randomized, double-blind, alendronate-controlled study to determine the efficacy and safety of romosozumab in the treatment of postmenopausal women with osteoporosis. UCB Data on File*, 2017

[21] Amgen. Study to determine the efficacy and safety of romosozumab in the treatment of postmenopausal women with osteoporosis (ARCH). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2012-2017 [cited 2021 Aug 04]. Available from: https://clinicaltrials.gov/ct2/show/NCT01631214. NLM Identifier: NCT01631214

[22] Amgen. Efficacy and safety of romosozumab treatment in postmenopausal women with osteoporosis (FRAME). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2012-2016 [cited 2021 Jul 01]. Available from: https://clinicaltrials.gov/ct2/show/NCT01575834. NLM Identifier: NCT01575834

[23] Miller PD, Hattersley G, Riis BJ, Williams GC, Lau E, Russo LA, et al. Effect of abaloparatide vs placebo on new vertebral fractures in postmenopausal women with osteoporosis: a randomized clinical trial. *JAMA* 2016;316(7):722-33.

[24] Chao M, Hua Q, Yingfeng Z, Guang W, Shufeng S, Yuzhen D, et al. Study on the role of zoledronic acid in treatment of postmenopausal osteoporosis women. *Pak J Med Sci* 2013;29(6):1381-4.

[25] Dursun N, Dursun E, Yalçin S. Comparison of alendronate, calcitonin and calcium treatments in postmenopausal osteoporosis. *Int J Clin Pract* 2001;55(8):505-9.

[26] Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA* 1998;280(24):2077-82.

[27] Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. *Lancet* 1996;348(9041):1535-41.

[28] Pols HA, Felsenberg D, Hanley DA, Stepán J, Muñoz-Torres M, Wilkin TJ, et al. Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in

postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. *Osteoporos Int* 1999;9(5):461-8.

[29] Hadji P, Zanchetta JR, Russo L, Recknor CP, Saag KG, McKiernan FE, et al. The effect of teriparatide compared with risedronate on reduction of back pain in postmenopausal women with osteoporotic vertebral fractures. *Osteoporos Int* 2012;23(8):2141-50.

[30] Liberman UA, Weiss SR, Bröll J, Minne HW, Quan H, Bell NH, et al. Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. *N Engl J Med* 1995;333(22):1437-43.

[31] Liu JL, Zhu HM, Huang QR, Zhang ZL, Li HL, Qin YJ, et al. Effects of raloxifene hydrochloride on bone mineral density, bone metabolism and serum lipids in Chinese postmenopausal women with osteoporosis: a multi-center, randomized, placebo-controlled clinical trial. *Chinese Medical Journal* 2004;117(7):1029-35.

[32] Lufkin EG, Whitaker MD, Nickelsen T, Argueta R, Caplan RH, Knickerbocker RK, et al. Treatment of established postmenopausal osteoporosis with raloxifene: a randomized trial. *J Bone Miner Res* 1998;13(11):1747-54.

[33] Delmas PD, Ensrud KE, Adachi JD, Harper KD, Sarkar S, Gennari C, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. *J Clin Endocrinol Metab* 2002;87(8):3609-17.

[34] Neer RM, Arnaud CD, Zanchetta JR, Prince R, Gaich GA, Reginster JY, et al. Effect of parathyroid hormone (1-34) on fractures and bone mineral density in postmenopausal women with osteoporosis. *N Engl J Med* 2001;344(19):1434-41.

[35] Hadji P, Gamerdinger D, Spieler W, Kann PH, Loeffler H, Articus K, et al. Rapid Onset and Sustained Efficacy (ROSE) study: results of a randomised, multicentre trial comparing the effect of zoledronic acid or alendronate on bone metabolism in postmenopausal women with low bone mass. *Osteoporos Int* 2012;23(2):625-33.

[36] Ensrud KE, Stock JL, Barrett-Connor E, Grady D, Mosca L, Khaw KT, et al. Effects of raloxifene on fracture risk in postmenopausal women: the Raloxifene Use for the Heart Trial. *J Bone Miner Res* 2008;23(1):112-20.

[37] Silverman SL, Christiansen C, Genant HK, Vukicevic S, Zanchetta JR, de Villiers TJ, et al. Efficacy of bazedoxifene in reducing new vertebral fracture risk in postmenopausal women with osteoporosis: results from a 3-year, randomized, placebo-, and active-controlled clinical trial. *J Bone Miner Res* 2008;23(12):1923-34.

[38] Kendler DL, Marin F, Zerbini CAF, Russo LA, Greenspan SL, Zikan V, et al. Effects of teriparatide and risedronate on new fractures in post-menopausal women with severe osteoporosis (VERO): a multicentre, double-blind, double-dummy, randomised controlled trial. *Lancet* 2018;391(10117):230-40.

[39] Reginster J, Minne HW, Sorensen OH, Hooper M, Roux C, Brandi ML, et al. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. *Osteoporos Int* 2000;11(1):83-91.

[40] Harris ST, Watts NB, Genant HK, McKeever CD, Hangartner T, Keller M, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal

osteoporosis: a randomized controlled trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. *JAMA* 1999;282(14):1344-52.

[41] Nakamura T, Fukunaga M, Nakano T, Kishimoto H, Ito M, Hagino H, et al. Efficacy and safety of once-yearly zoledronic acid in Japanese patients with primary osteoporosis: two-year results from a randomized placebo-controlled double-blind study (ZOledroNate treatment in Efficacy to osteoporosis; ZONE study). *Osteoporos Int* 2017;28(1):389-98.

[42] Adami S, Passeri M, Ortolani S, Broggini M, Carratelli L, Caruso I, et al. Effects of oral alendronate and intranasal salmon calcitonin on bone mass and biochemical markers of bone turnover in postmenopausal women with osteoporosis. *Bone* 1995;17(4):383-90.

[43] Adami S, San Martin J, Munoz-Torres M, Econs MJ, Xie L, Dalsky GP, et al. Effect of raloxifene after recombinant teriparatide [hPTH(1-34)] treatment in postmenopausal women with osteoporosis. *Osteoporos Int* 2008;19(1):87-94.

[44] Aki S, Gunay Gulbabab R, Eskiyurta N. Effect of alendronate on bone density and bone markers in postmenopausal osteoporosis. *J Back Musculoskelet Rehabil* 2004;17(2003/2004):27-31.

[45] McClung MR, Lewiecki EM, Cohen SB, Bolognese MA, Woodson GC, Moffett AH, et al. Denosumab in postmenopausal women with low bone mineral density. *N Engl J Med* 2006;354(8):821-31.

[46] Tsai JN, Uihlein AV, Lee H, Kumbhani R, Siwila-Sackman E, McKay EA, et al. Teriparatide and denosumab, alone or combined, in women with postmenopausal osteoporosis: the DATA study randomised trial. *Lancet* 2013;382(9886):50-6.

[47] Brown JP, Prince RL, Deal C, Recker RR, Kiel DP, de Gregorio LH, et al. Comparison of the effect of denosumab and alendronate on BMD and biochemical markers of bone turnover in postmenopausal women with low bone mass: a randomized, blinded, phase 3 trial. *J Bone Miner Res* 2009;24(1):153-61.

[48] Bone HG, Bolognese MA, Yuen CK, Kendler DL, Wang H, Liu Y, et al. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. *J Clin Endocrinol Metab* 2008;93(6):2149-57.

[49] Luckey M, Kagan R, Greenspan S, Bone H, Kiel RD, Simon J, et al. Once-weekly alendronate 70 mg and raloxifene 60 mg daily in the treatment of postmenopausal osteoporosis. *Menopause* 2004;11(4):405-15.

[50] Sambrook PN, Geusens P, Ribot C, Solimano JA, Ferrer-Barriendos J, Gaines K, et al. Alendronate produces greater effects than raloxifene on bone density and bone turnover in postmenopausal women with low bone density: results of EFFECT (Efficacy of FOSAMAX versus EVISTA Comparison Trial) International. *J Intern Med* 2004;255(4):503-11.

[51] Eastell R, Nickelsen T, Marin F, Barker C, Hadji P, Farrerons J, et al. Sequential treatment of severe postmenopausal osteoporosis after teriparatide: final results of the randomized, controlled European Study of Forsteo (EUROFORS). *J Bone Miner Res* 2009;24(4):726-36.

[52] Rosen CJ, Hochberg MC, Bonnick SL, McClung M, Miller P, Broy S, et al. Treatment with once-weekly alendronate 70 mg compared with once-weekly risedronate 35 mg in women with postmenopausal osteoporosis: a randomized double-blind study. *J Bone Miner Res* 2005;20(1):141-51.

[53] Reid DM, Hosking D, Kendler D, Brandi ML, Wark JD, Marques-Neto JF, et al. A comparison of the effect of alendronate and risedronate on bone mineral density in postmenopausal women with osteoporosis: 24-month results from FACTS-International. *Int J Clin Pract* 2008;62(4):575-84.

[54] Fogelman I, Ribot C, Smith R, Ethgen D, Sod E, Reginster JY. Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab* 2000;85(5):1895-900.

[55] Grey A, Bolland M, Wattie D, Horne A, Gamble G, Reid IR. Prolonged antiresorptive activity of zoledronate: a randomized, controlled trial. *J Bone Miner Res* 2010;25(10):2251-5.

[56] Johnell O, Scheele WH, Lu Y, Reginster JY, Need AG, Seeman E. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. *J Clin Endocrinol Metab* 2002;87(3):985-92.

[57] McClung MR, Bolognese MA, Sedarati F, Recker RR, Miller PD. Efficacy and safety of monthly oral ibandronate in the prevention of postmenopausal bone loss. *Bone* 2009;44(3):418-22.

[58] McClung MR, Grauer A, Boonen S, Bolognese MA, Brown JP, Diez-Perez A, et al. Romosozumab in postmenopausal women with low bone mineral density. *N Engl J Med* 2014;370(5):412-20.

[59] Amgen Inc. A study to evaluate the safety and efficacy of denosumab compared with zoledronic acid in postmenopausal women with osteoporosis previously treated with oral bisphosphonates. EUCTR2012-001821-28-ES. In: WHO International Clinical Trials Registry Platform (ICTRP) [Internet]. Geneva: World Health Organization (WHO), 2012 [accessed 4.8.21].

[60] Miller PD, Epstein S, Sedarati F, Reginster JY. Once-monthly oral ibandronate compared with weekly oral alendronate in postmenopausal osteoporosis: results from the head-to-head MOTION study. *Curr Med Res Opin* 2008;24(1):207-13.

[61] Valimaki MJ, Farrerons-Minguella J, Halse J, Kroger H, Maroni M, Mulder H, et al. Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. *Clin Ther* 2007;29(9):1937-49.

[62] Bonnick S, Broy S, Kaiser F, Teutsch C, Rosenberg E, DeLucca P, et al. Treatment with alendronate plus calcium, alendronate alone, or calcium alone for postmenopausal low bone mineral density. *Curr Med Res Opin* 2007;23(6):1341-9.

[63] Eastell R, Nagase S, Ohyama M, Small M, Sawyer J, Boonen S, et al. Safety and efficacy of the cathepsin K inhibitor ONO-5334 in postmenopausal osteoporosis: the OCEAN study. *J Bone Miner Res* 2011;26(6):1303-12.

[64] Recknor C, Czerwinski E, Bone HG, Bonnick SL, Binkley N, Palacios S, et al. Denosumab compared with ibandronate in postmenopausal women previously treated with bisphosphonate therapy: a randomized open-label trial. *Obstet Gynecol* 2013;121(6):1291-9.

[65] Reid IR, Bolland M, Wong S, et al. A randomized controlled trial of annual low dose zoledronate in postmenopausal women. *Osteoporos Int* 2011;2:S526.

[66] Roux C, Hofbauer LC, Ho PR, Wark JD, Zillikens MC, Fahrleitner-Pammer A, et al. Denosumab compared with risedronate in postmenopausal women suboptimally adherent to alendronate therapy: efficacy and safety results from a randomized open-label study. *Bone* 2014;58:48-54.

[67] Bock O, Borst H, Beller G, Armbrecht G, Degner C, Martus P, et al. Impact of oral ibandronate 150 mg once monthly on bone structure and density in post-menopausal osteoporosis or osteopenia derived from in vivo muCT. *Bone* 2012;50(1):317-24.

[68] Kendler DL, Roux C, Benhamou CL, Brown JP, Lillestol M, Siddhanti S, et al. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women transitioning from alendronate therapy. *J Bone Miner Res* 2010;25(1):72-81.

[69] Tan W, Sun J, Zhou L, Li Y, Wu X. Randomized trial comparing efficacies of zoledronate and alendronate for improving bone mineral density and inhibiting bone remodelling in women with post-menopausal osteoporosis. *J Clin Pharm Ther* 2016;41(5):519-23.

[70] Tucci JR, Tonino RP, Emkey RD, Peverly CA, Kher U, Santora AC, 2nd. Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am J Med* 1996;101(5):488-501.

[71] Um MJ, Cho EA, Jung H. Combination therapy of raloxifene and alendronate for treatment of osteoporosis in elderly women. *J Menopausal Med* 2017;23(1):56-62.

[72] Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. *Syst Rev* 2021;10(1):39.

[73] Söreskog E, Lindberg I, Kanis JA, Åkesson KE, Willems D, Lorentzon M, et al. Cost-effectiveness of romosozumab for the treatment of postmenopausal women with severe osteoporosis at high risk of fracture in Sweden. *Osteoporos Int* 2021;32(3):585-94.

[74] Soreskog E, Borgstrom F, Lindberg I, Strom O, Willems D, Libanati C, et al. A novel economic framework to assess the cost-effectiveness of bone-forming agents in the prevention of fractures in patients with osteoporosis. *Osteoporos Int* 2021;32(7):1301-1311.

[75] Quantify. *Quantify Research Data on File. Swedish Osteoporosis Research Database*, 2017

[76] National Osteoporosis Guidelines Group. *NOGG 2017: Clinical guideline for the prevention and treatment of osteoporosis [Internet]*, 2017 [accessed 1.7.21] Available from: https://www.sheffield.ac.uk/NOGG/NOGG%20Guideline%202017.pdf

[77] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 (2018 revision) [Internet]*. London: NICE, 2013 (revised 2018 [accessed 1.7.21] Available from: https://www.nice.org.uk/process/pmg9/

[78] Hiligsmann M, Reginster JY, Tosteson ANA, Bukata SV, Saag KG, Gold DT, et al. Recommendations for the conduct of economic evaluations in osteoporosis: outcomes of an experts' consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the US branch of the International Osteoporosis Foundation. *Osteoporos Int* 2019;30(1):45-57.

[79] Svedbom A, Hernlund E, Ivergard M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: a compendium of country-specific reports. *Arch Osteoporos* 2013;8:137.

[80] Hernlund E, Svedbom A, Ivergard M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch Osteoporos* 2013;8:136.

[81] Singer BR, McLauchlan GJ, Robinson CM, Christie J. Epidemiology of fractures in 15,000 adults: the influence of age and gender. *J Bone Joint Surg Br* 1998;80(2):243-8.

[82] van der Velde RY, Wyers CE, Curtis EM, Geusens P, van den Bergh JPW, de Vries F, et al. Secular trends in fracture incidence in the UK between 1990 and 2012. *Osteoporos Int* 2016;27(11):3197-3206.

[83] Kanis JA, Johnell O, Oden A, Sembo I, Redlund-Johnell I, Dawson A, et al. Long-term risk of osteoporotic fracture in Malmö. *Osteoporos Int* 2000;11(8):669-74.

[84] Kanis JA, Johnell O, Oden A, De Laet C, Oglesby A, Jönsson B. Intervention thresholds for osteoporosis. *Bone* 2002;31(1):26-31.

[85] Soreskog E, Strom O, Spangeus A, Akesson KE, Borgstrom F, Banefelt J, et al. Risk of major osteoporotic fracture after first, second and third fracture in Swedish women aged 50years and older. *Bone* 2020;134:115286.

[86] Banefelt J, Akesson KE, Spangeus A, Ljunggren O, Karlsson L, Strom O, et al. Risk of imminent fracture following a previous fracture in a Swedish database study. *Osteoporos Int* 2019;30(3):601-9.

[87] National Institute for Health and Care Excellence. *Assessment Group Report: Non-bisphosphonates for treating osteoporosis [ID901]*. Sheffield: School of Health and Related Research (ScHARR), University of Sheffield, 2019

[88] Simpson EL, Martyn-St James M, Hamilton J, Wong R, Gittoes N, Selby P, et al. Clinical effectiveness of denosumab, raloxifene, romosozumab, and teriparatide for the prevention of osteoporotic fragility fractures: a systematic review and network meta-analysis. *Bone* 2020;130:115081.

[89] Willems D, Javaid MK, Pinedo-Villaneuva R, Libanati C, Yehoshua A, Charokopou M, et al. UCB Data on File - SLR NMA. The importance of time-point specific indirect treatment comparisons of osteoporosis treatments: a systematic literature review (SLR) and network meta-analysis (NMA). *Clin Ther* 2021.

[90] Kanis JA, Oden A, Johnell O, Jonsson B, de Laet C, Dawson A. The burden of osteoporotic fractures: a method for setting intervention thresholds. *Osteoporos Int* 2001;12(5):417-27.

[91] Li L, Roddam A, Gitlin M, Taylor A, Shepherd S, Jick S. Concurrent Oral 8 – Innovations [OP54–OP58]: OP54. Non-persistence to anti-osteoporosis medications in the UK using the General Practice Research Database (GPRD). *Rheumatology (Oxford)* 2010;49(Suppl 1):i23-i25.

[92] Solomon DH, Avorn J, Katz JN, Finkelstein JS, Arnold M, Polinski JM, et al. Compliance with osteoporosis medications. *Arch Intern Med* 2005;165(20):2414-9.

[93] Huybrechts KF, Ishak KJ, Caro JJ. Assessment of compliance with osteoporosis treatment and its consequences in a managed care population. *Bone* 2006;38(6):922-8.

[94] Borgström F, Johnell O, Kanis JA, Oden A, Sykes D, Jönsson B. Cost effectiveness of raloxifene in the treatment of osteoporosis in Sweden: an economic evaluation based on the MORE study. *Pharmacoeconomics* 2004;22(17):1153-65.

[95] Davis S, Martyn-St James M, Sanderson J, Stevens J, Goka E, Rawdin A, et al. *Bisphosphonates for preventing osteoporotic fragility fractures (including a partial update of NICE technology appraisal guidance 160 and 161). Technology assessment report: final report to the National Institute for Health and Care Excellence [Internet]*. London: National Institute for Health and Care Excellence, 2015

[accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta464/documents/osteoporosis-prevention-bisphosphonates-inc-part-rev-ta160-ta161-id782-assessment-report2

[96] Li L, Roddam A, Gitlin M, Taylor A, Shepherd S, Shearer A, et al. Persistence with osteoporosis medications among postmenopausal women in the UK General Practice Research Database. *Menopause* 2012;19(1):33-40.

[97] Karlsson L, Lundkvist J, Psachoulia E, Intorcia M, Strom O. Persistence with denosumab and persistence with oral bisphosphonates for the treatment of postmenopausal osteoporosis: a retrospective, observational study, and a meta-analysis. *Osteoporos Int* 2015;26(10):2401-11.

[98] McHorney CA, Spain CV. Frequency of and reasons for medication non-fulfillment and non-persistence among American adults with chronic disease in 2008. *Health Expect* 2011;14(3):307-20.

[99] National Institute for Health and Care Excellence. *Osteoporosis: assessing the risk of fragility fracture. NICE clinical guideline 146 [Internet]*. London: National Institute for Health and Care Excellence, 2012 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/cg146

[100] Morley J, Moayyeri A, Ali L, Taylor A, Feudjo-Tepie M, Hamilton L, et al. Persistence and compliance with osteoporosis therapies among postmenopausal women in the UK Clinical Practice Research Datalink. *Osteoporos Int* 2020;31(3):533-45.

[101] Vanelli M, Pedan A, Liu N, Hoar J, Messier D, Kiarsis K. The role of patient inexperience in medication discontinuation: a retrospective analysis of medication nonpersistence in seven chronic illnesses. *Clin Ther* 2009;31(11):2628-52.

[102] Kanis JA, McCloskey EV, Jonsson B, Cooper A, Ström O, Borgström F. An evaluation of the NICE guidance for the prevention of osteoporotic fragility fractures in postmenopausal women. *Arch Osteoporos* 2010;5(1-2):19-48.

[103] Jonsson B, Strom O, Eisman JA, Papaioannou A, Siris ES, Tosteson A, et al. Cost-effectiveness of Denosumab for the treatment of postmenopausal osteoporosis. *Osteoporos Int* 2011;22(3):967-82.

[104] Tandvårds- och läkemedelsförmånsverket (The Dental and Pharmaceutical Benefits Agency). *Underlag för beslut om subvention - Nyansökan [Decision for reimbursement]*, 2020

[105] Black DM, Schwartz AV, Ensrud KE, Cauley JA, Levis S, Quandt SA, et al. Effects of continuing or stopping alendronate after 5 years of treatment: the Fracture Intervention Trial Long-term Extension (FLEX): a randomized trial. *JAMA* 2006;296(24):2927-38.

[106] Black DM, Reid IR, Boonen S, Bucci-Rechtweg C, Cauley JA, Cosman F, et al. The effect of 3 versus 6 years of zoledronic acid treatment of osteoporosis: a randomized extension to the HORIZON-Pivotal Fracture Trial (PFT). *J Bone Miner Res* 2012;27(12):243-54.

[107] Lindsay R, Scheele WH, Neer R, Pohl G, Adami S, Mautalen C, et al. Sustained vertebral fracture risk reduction after withdrawal of teriparatide in postmenopausal women with osteoporosis. *Arch Intern Med* 2004;164(18):2024-30.

[108] Davis S, Martyn-St James M, Sanderson J, Stevens J, Goka E, Rawdin A, et al. A systematic review and economic evaluation of bisphosphonates for the prevention of fragility fractures. *Health Technol Assess* 2016;20(78):1-406.

[109] Strom O, Landfeldt E, Garellick G. Residual effect after oral bisphosphonate treatment and healthy adherer effects--the Swedish Adherence Register Analysis (SARA). *Osteoporos Int* 2015;26(1):315-25.

[110] Tsourdi E, Langdahl B, Cohen-Solal M, Aubry-Rozier B, Eriksen EF, Guañabens N, et al. Discontinuation of denosumab therapy for osteoporosis: a systematic review and position statement by ECTS. *Bone* 2017;105:11-7.

[111] Miller PD, Bolognese MA, Lewiecki EM, McClung MR, Ding B, Austin M, et al. Effect of denosumab on bone density and turnover in postmenopausal women with low bone mass after long-term continued, discontinued, and restarting of therapy: a randomized blinded phase 2 clinical trial. *Bone* 2008;43(2):222-9.

[112] Davis S, Simpson E, Hamilton J, James MM, Rawdin A, Wong R, et al. Denosumab, raloxifene, romosozumab and teriparatide to prevent osteoporotic fragility fractures: a systematic review and economic evaluation. *Health Technology Assessment* 2020;24(29):1-314.

[113] Office for National Statistics (ONS). *National life tables, UK: 2012 to 2014 [Internet]*, 2020 [accessed 1.7.21] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables

[114] Strom O, Borgstrom F, Sen SS, Boonen S, Haentjens P, Johnell O, et al. Cost-effectiveness of alendronate in the treatment of postmenopausal women in 9 European countries--an economic evaluation based on the fracture intervention trial. *Osteoporos Int* 2007;18(8):1047-61.

[115] Office for National Statistics (ONS). *National life tables, UK: 2017 to 2019 [Internet]*, 2020 [accessed 28.9.21] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables

[116] Johnell O, Kanis JA, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Mortality after osteoporotic fractures. *Osteoporos Int* 2004;15(1):38-42.

[117] Barrett JA, Baron JA, Beach ML. Mortality and pulmonary embolism after fracture in the elderly. *Osteoporos Int* 2003;14(11):889-94.

[118] Oden A, Dawson A, Dere W, Johnell O, Jonsson B, Kanis JA. Lifetime risk of hip fractures is underestimated. *Osteoporos Int* 1998;8(6):599-603.

[119] Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B, Oglesby AK. The components of excess mortality after hip fracture. *Bone* 2003;32(5):468-473.

[120] Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone* 2004;35(2):375-82.

[121] National Institute for Health and Care Excellence. *Raloxifene and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE technology appraisal guidance 161 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta161

[122] National Institute for Health and Care Excellence. *Raloxifene for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE technology appraisal guidance 160 [Internet]*. London: National Institute for Health and Care Excellence, 2008 [accessed 1.7.21] Available from: https://www.nice.org.uk/guidance/ta160

[123] National Institute for Health and Care Excellence. *Romosozumab for treating severe osteoporosis [ID3936]. Clarification letter. Single technology appraisal*. London: NICE, 2021

[124] Szende A, Janssen B, Cabases J. *Self-reported population health: an international perspective based on EQ-5D*. The Netherlands: Springer, 2014.

[125] Scottish Medicines Consortium. *Medicines advice romosozumab (Evenity)*, 2020

[126] Kanis JA, Johansson H, Oden A, Harvey NC, Gudnason V, Sanders KM, et al. Characteristics of recurrent fractures. *Osteoporos Int* 2018;29(8):1747-57.

[127] Adachi JD, Lyles KW, Colón-Emeric CS, Boonen S, Pieper CF, Mautalen C, et al. Zoledronic acid results in better health-related quality of life following hip fracture: the HORIZON-Recurrent Fracture Trial. *Osteoporos Int* 2011;22(9):2539-49.

[128] Blomfeldt R, Törnkvist H, Ponzer S, Söderqvist A, Tidermark J. Comparison of internal fixation with total hip replacement for displaced femoral neck fractures. Randomized, controlled trial performed at four years. *J Bone Joint Surg Am* 2005;87(8):1680-8.

[129] Ekström W, Németh G, Samnegård E, Dalen N, Tidermark J. Quality of life after a subtrochanteric fracture: a prospective cohort study on 87 elderly patients. *Injury* 2009;40(4):371-6.

[130] Lindgren P, Lindström J, Tuomilehto J, Uusitupa M, Peltonen M, Jönsson B, et al. Lifestyle intervention to prevent diabetes in men and women with impaired glucose tolerance is cost-effective. *Int J Technol Assess Health Care* 2007;23(2):177-83.

[131] Joint Formulary Committee. *British National Formulary [Internet]*. London: BMJ Group and Pharmaceutical Press, 2021 [accessed 1.7.21]. Available from: https://bnf.nice.org.uk/

[132] Eli Lilly and Company Limited. Forsteo 20 micrograms/80 microlitres solution for injection in pre-filled pen (teriparatide). Summary of product characteristics [Internet]. electronic Medicines Compendium (eMC), 2020 [accessed 1.7.21]. Available from: https://www.medicines.org.uk/emc/product/2215

[133] Curtis LA, Burns A. *Unit costs of health and social care 2020 [Internet]*. Canterbury: Personal Social Services Research Unit (PSSRU), University of Kent, 2020 [accessed 1.7.21] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

[134] NHS. *2020/21 National Tariff Payment System. The National Tariff Workbook 2020/21 [Internet]*: National Health Service, 2020 [accessed 1.7.21] Available from: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/

[135] Borgstrom F, Jonsson B, Strom O, Kanis JA. 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):1781-93.

[136] Gutierrez L, Roskell N, Castellsague J, Beard S, Rycroft C, Abeysinghe S, et al. Clinical burden and incremental cost of fractures in postmenopausal women in the United Kingdom. *Bone* 2012;51(3):324-31.

[137] Gutierrez L, Roskell N, Castellsague J, Beard S, Rycroft C, Abeysinghe S, et al. Study of the incremental cost and clinical burden of hip fractures in postmenopausal women in the United Kingdom. *J Med Econ* 2011;14(1):99-107.

[138] Office for National Statistics (ONS). Consumer price inflation tables [Internet]. ONS, 2021 [accessed 4.8.21]. Available from: https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation

[139] Ryder S, Fox K, Rane P, Armstrong N, Wei CY, Deshpande S, et al. A Systematic Review of Direct Cardiovascular Event Costs: An International Perspective. *Pharmacoeconomics* 2019;37(7):895-919.

[140] Danese MD, Gleeson M, Kutikova L, Griffiths RI, Azough A, Khunti K, et al. Estimating the economic burden of cardiovascular events in patients receiving lipid-modifying therapy in the UK. *BMJ OPEN* 2016;6(8):e011805.

[141] Chapman RH, Liu LZ, Girase PG, Straka RJ. Determining initial and follow-up costs of cardiovascular events in a US managed care population. *BMC Cardiovasc Disord* 2011;11:11.

[142] van Staa TP, Kanis JA, Geusens P, Boonen A, Leufkens HG, Cooper C. The cost-effectiveness of bisphosphonates in postmenopausal women based on individual long-term fracture risks. *Value Health* 2007;10(5):348-57.

[143] UCB Pharma Ltd. *UCB Data on File - National Institute of Health and Care Excellence scientific advice. Preliminary independent model advice + clarification questions. Cost-effectiveness model for romosozumab for the treatment of osteoporosis*, 2017

[144] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Romosozumab for treating severe osteoporosis [ID3936]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and ERG report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the ERG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on 15 October 2021** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and separately highlight information that is submitted as '[redacted]' in turquoise, all information submitted as '[redacted]' in yellow, and all information submitted as '[redacted]' in pink.

Issue 1 Unclear Description of ERG preferred estimates of persistence

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 4.13, Page 109 The ERG report notes "Therefore, the ERG prefers to inform persistence with alendronate, regardless of whether it is given as a standalone treatment or after romosozumab, using the same study" It appears that the ERG has chosen to use persistence values for naïve patients from Morley et al. 2020 to inform the persistence rates for alendronate alone,[1] but persistence rates from non-naïve patients to inform persistence for alendronate after romosozumab.	Please could the ERG report provide additional explanation in this table, and the associated text, to explain how the persistence estimates for alendronate alone, and alendronate after romosozumab have been derived from the Morley et al. 2020 paper. – currently this is not clear to the reader, Please could the ERG provide justification for the derivation of these two sets of persistence estimates.	Table 4.13 is not clear about how the two sources of persistence to alendronate have been derived from Morley et al. 2020 – it is important that the ERG provides additional explanation so that this is clear to the reader. Further, it is important that the ERG provides additional justification about why it believes Morley et al. 2020 is the most appropriate source of persistence estimates, and why it should be considered clinically plausible for the persistence to alendronate after romosozumab to be less than half of the persistence to alendronate alone. While Morley et al. 2020 provides more up-to-date estimates of persistence in general UK population, several limitations of the study for application to this analysis should be considered.[1] Firstly, Morley et al. looked into general users of any osteoporosis medication irrespective of their diagnosis or fracture status. As indicated in Table 1, only 35.6% of patients receiving oral	The ERG preferred to source persistence rates from Morley et al. 2020 because it is the most recent source of persistence estimates based on UK real-world evidence. The ERG has made the following amendment on p. 109 ("…and used…"): "The ERG identified a more recent study by Morley et al. 2020 on persistence with osteoporosis therapies that also made use of UK CPRD data.100 The ERG preferred to use this more recent source of persistence estimates for their base-case, and used the estimates for non-naïve patients for alendronate after romosozumab and the estimates from naïve patients for alendronate alone." The ERG has further included two footnotes in Table 4.13 to indicate having used persistence values for naïve patients for alendronate alone, and persistence rates from non-naïve patients for

		bisphosphonates had a history of fracture and 29.8% of them were diagnosed by osteoporosis. These numbers were about 5% and 10%, respectively, higher in naïve-treated patients (supplementary Table S1), which indicates that naïve patients are presented with more severe osteoporosis. However, it should be noted that even the naïve patient population is less likely to represent the imminent fracture risk patients that are going to be target population for romosozumab (all having a major osteoporotic fracture in the last 2 years). Hence, one can argue that even the naïve measures from this study can be under-estimates for actual alendronate persistence after romosozumab in the target population. Another point to consider is the timeline of study by Morley et al. (2010 to 2015). This was the early post-FRAX era and there were many uncertainties in management of patients with osteoporosis. Many physicians would discontinue treatment for their patients (or start drug holiday periods) if patients were not considered high risk by	alendronate after romosozumab. The company indicates several limitations of the study by Morley et al. but it is not clear whether these or similar limitations apply to their preferred source of Li et al. The company should provide justification for why they think that Li et al. is a more appropriate source to inform persistence estimates than the study by Morley et al. which effectively is an update of the study by Li et al. This is because both studies address the same research question using UK CPRD data, with the difference being that Morley et al. used more recent data. As such, it is likely that the same limitations that the company notes for Morley et al. also apply to Li et al. There is currently no evidence to support the assumptions regarding the impact of the PSP. Once the PSP is in place, additional evidence could be collected by the company on real-world persistence with romosozumab (and subsequent alendronate). The ERG therefore preferred

		FRAX as recommended by the National Osteoporosis Guideline Group (NOGG).[2] Therefore, the overall persistence can be under-estimated in this study. Another important factor to consider is the impact of patient support programme (PSP) and clinical interactions on the persistence of treatment. The ERG argues that there is no evidence for this claim, while there are numerous supporting studies both in the field of osteoporosis (for teriparatide;[3] and other chronic diseases[4-6]). This is very important in the case of osteoporosis, which is a silent disease until it manifests itself with a fracture. The Company believe patients who are going to actively receive 12 doses of romosozumab with clinical interactions and a PSP making them aware of their disease will be more persistent in taking their follow-up drug of alendronate for the period of their treatment. Hence, the measures currently used by the ERG are highly under-estimating the actual persistence level of treatment in the target population of romosozumab.	to assume a situation where the PSP is not in place for their base-case.

Issue 2 Clinically implausible persistence to romosozumab

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Table 4.13, Page 109 The ERG notes that it's preferred assumption is 80% romosozumab persistence at Month 6, based on 80% romosozumab persistence at Month 12.	Please could the ERG update the persistence to romosozumab at Month 6 – the current estimate of 80% appears to be factually inaccurate, and the persistence to romosozumab estimate should be increased to a higher estimate that lies between 100% and 80%.	Based on the published assumption of 80% persistence to romosozumab at Month 12, it is not clinically plausible to assume 80% persistence to romosozumab at Month 6. It is clinically implausible to assume that 20% of patients discontinue romosozumab during the first six months of treatment, but then 0% of patients discontinue treatment with romosozumab during the second six months of treatment. Accordingly, please could the ERG update this estimate if it is factually inaccurate, or provide additional justification for why the ERG believes that this assumption is clinically plausible. The Company believes that the persistence to romosozumab should be increased to a higher estimate that lies between 100% and 80%. The ERG should additional consider clinical plausibility, and provide further justifications, across all scenarios where they have considered alternative estimates for persistence.	The ERG preferred using lower estimates than the 90% persistence rate that was used by the company based on what was observed in ARCH, to account for real-world persistence likely being lower than in the context of a clinical trial. Instead of using [redacted] at both 6 months and 12 months as assumed by the company, the ERG assumed 80% for both time points. The ERG notes that the assumption that persistence is the same at both time points was also made by the company and would represent a situation where all treatment discontinuation takes place in the first 6 months with no additional discontinuation between months 6 and 12. Therefore, this is not clinically implausible.

Issue 3 Incorrect description of Company NMAs

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 70 The ERG reports states *"In total, 11 NMAs were presented by the company, covering five distinct outcomes at three timepoints."*	*"In total, ~~11~~ **12** NMAs were presented by the company, covering ~~five~~ **three fracture outcomes at three timepoints, as well as three BMD outcomes which were not timepoint specific.***	The Company conducted 12 NMAs for six distinct outcomes including: • New vertebral fracture (Month 12, 24 and 36) • Non-vertebral fracture (Month 12, 24 and 36) • Hip fracture (Month 12, 24 and 36) • BMD (Total Hip) • BMD (Femoral Neck) • BMD (Lumbar Spine)	The text was amended accordingly.

Issue 4 Incorrect description of comparison between zoledronate and placebo

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 81 The ERG report notes: Zoledronate: There was no observed heterogeneity between the comparison of alendronate and placebo	Zoledronate: There was no observed heterogeneity between the comparison of ~~alendronate~~ **zoledronate** and placebo	This section discusses the comparison between zoledronate and placebo, and mention of alendronate should be updated to correctly mention zoledronate.	Changed as suggested.

Issue 5 Inclusion of Fracture Costs in the ERG's Scenarios

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 125, The ERG report notes: "The first-year costs of hip, vertebral and NHNV fractures that were sourced from Gutiérrez et al., were based on the total costs…"	"The first-year costs of hip, vertebral and NHNV fractures that were sourced from Gutiérrez et al., were based on the ~~total~~ **incremental** costs…" Please could the ERG re-run the relevant scenarios using the corrected fracture cost calculations, and including rehabilitation costs, to avoid substantial underestimation of the costs associated with fractures.	The ERG notes that the Company has not used the fracture costs adjusted for controls – this statement is incorrect; the Company used the controlled costs in their base case analysis. The Company also included rehabilitation costs in the base case analysis. These are an important cost associated with fractures, and the ERG's exclusion of rehabilitation costs substantially underestimates the costs associated with fractures. For example, Gutierrez et al. 2011 calculated the rehabilitation costs separately, and concluded that the total cost of hip fracture as £9,936 (£8,752 adjusted for controls) in the 2006/2007 cost year. Consequently, the ERG's preferred cost for hip fractures of £5,369 appears to be factually inaccurate based on Gutierrez et al – please could the ERG provide additional justification for this assumption and/or re-run these analyses using the corrected calculations.	In the request for clarification, the ERG asked the company (Question B21) to "please confirm that the total (i.e. not the incremental) cost estimates from Gutiérrez et al. were used in the analysis for patients who had a fracture but not for those who did not have a fracture, and please justify the appropriateness of this approach." The company responded: "The total costs based on Gutiérrez et al. (2011) were used in the analysis for patients who suffered a fracture. No medical costs were applied for those who did not suffer fracture. The total costs rather than incremental costs were used in the model for two reasons. Firstly, both the incremental and total cost in the Gutiérrez et al. (2011) study are likely underestimated due to censoring bias. The follow-up time is shorter for the fractured cohort compared with the non-

			fractured cohort, which is likely due to higher mortality in the fractured cohort. This is not adjusted for in the two source papers. This underestimates costs but it is unknown to what extent. In the model, cost is applied for each cycle after the fracture (until the patient dies) and, as the cost input is unadjusted for censoring, the total costs would be underestimated in the model as well. Secondly, using total costs as opposed to incremental costs is the standard in economic evaluations, for example, in Jönsson, et al. (2011).” As such, it appears to be factually incorrect that incremental cost estimates rather than total costs were used. However, the ERG has performed a scenario analysis using the same cost estimates as the company used for their base-case, which had only a minor impact on the cost effectiveness results.

Issue 6 Insufficient Description of the ERG's Scenario 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 147 The ERG report notes *"the ERG performed a set of scenario analyses where the "imminent risk" of fracture was removed from the analysis"*	Please could the ERG provide additional description about the technical methodology used to remove the "imminent risk of fracture" from the analyses?	The Company was unable to replicate the ERG's scenario analyses where imminent risk of fracture was removed from the model. Please could the ERG provide additional details on how this was implemented, so that the Company can fully understand and replicate the assumptions used in these scenarios.	The scenarios where the imminent risk of fracture was removed were conducted by selecting "No" in the dropdown list in "Main settings" cell I13 ("Enable recent fracture risk estimation"). This approach was in line with the approach used by the company in their response to clarification question B3.

Confidentiality Highlighting

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG response
Page 40, Table 3.9 The ERG report presents BMD Outcomes: N, LS Mean (SE) in Table 3.9. The N values should be highlighted yellow (AIC) as these are not available in the public domain.	The ERG report presents BMD Outcomes: N, LS Mean (SE) in Table 3.9. The N values should be highlighted yellow (AIC) as these are not available in the public domain. he N values should be highlighted yellow (AIC) as these are not available in the public domain.	Please [redacted] the following values as AIC in Table 3.9: • BMD at the lumbar spine at 12 months: [redacted] • BMD at the lumbar spine at 24 months: [redacted] • BMD at the lumbar spine at 36 months: [redacted]	Changed as suggested. Please note that the corresponding values reported for romoso-zumab were not changed.

		• BMD at the total hip at 12 months: [redacted] • BMD at the total hip at 24 months: [redacted] • BMD at the total hip at 36 months: [redacted] • BMD at the femoral neck at 12 months: [redacted] • BMD at the femoral neck at 24 months: [redacted] • BMD at the femoral neck at 36 months: [redacted]	

References

1. Morley J, Moayyeri A, Ali L, et al. Persistence and compliance with osteoporosis therapies among postmenopausal women in the UK Clinical Practice Research Datalink. Osteoporosis International 2020;31:533-545.
2. Compston J, Bowring C, Cooper A, et al. Diagnosis and management of osteoporosis in postmenopausal women and older men in the UK: National Osteoporosis Guideline Group (NOGG) update 2013. Maturitas 2013;75:392-6.
3. Sato M, Tsujimoto M, Kajimoto K, et al. Effect of a patient-support program on once-daily teriparatide adherence and persistence in the Japan Fracture Observational Study (JFOS). Arch Osteoporos 2018;13:74.
4. Srulovici E, Garg V, Ghilai A, et al. Is Patient Support Program Participation Associated with Longer Persistence and Improved Adherence Among New Users of Adalimumab? A Retrospective Cohort Study. Adv Ther 2018;35:655-665.
5. Zhou FL, Yeaw J, Karkare SU, et al. Impact of a structured patient support program on adherence and persistence in basal insulin therapy for type 2 diabetes. BMJ Open Diabetes Res Care 2018;6:e000593.
6. Chang A, Stokes J, Priestman L, et al. Impact of a Patient Support Program on Patient Beliefs About Neovascular Age-Related Macular Degeneration and Persistence to Anti-Vascular Endothelial Growth Factor Therapy. Patient preference and adherence 2021;15:511-521.