Systematic review of the use of bone turnover markers for monitoring the response to osteoporosis treatment: the secondary prevention of fractures, and primary prevention of fractures in high-risk groups

Jane Burch, Stephen Rice, Huiqin Yang, Aileen Neilson, Lisa Stirk, Roger Francis, Paul Holloway, Peter Selby and Dawn Craig
Systematic review of the use of bone turnover markers for monitoring the response to osteoporosis treatment: the secondary prevention of fractures, and primary prevention of fractures in high-risk groups

Jane Burch,¹ Stephen Rice,¹ Huiqin Yang,¹ Aileen Neilson,¹ Lisa Stirk,¹ Roger Francis,² Paul Holloway,³ Peter Selby⁴ and Dawn Craig¹*

¹Centre for Reviews and Dissemination, York, UK
²Institute for Ageing and Health, Newcastle University, Newcastle upon Tyne, UK
³Imperial College Healthcare Trust, St Mary’s Hospital, London, UK
⁴Department of Medicine, Manchester Royal Infirmary, Manchester, UK

*Corresponding author

Declared competing interests of authors: Roger Francis has served on an advisory board for the manufacturers of denosumab, Amgen/GlaxoSmithKline (GSK); served as an advisor on a study of strontium ranelate to Servier; received fees as an expert advisor in a legal case regarding a potential bisphosphonate patient case; and received lecture fees from Servier, Amgen/GSK and Shire Pharmaceuticals. Roger Francis has no competing interests related to the bone marker tests being evaluated. Paul Holloway is director of a bone biochemical marker diagnostic supra-regional analytical and advisory service that supports his hospital clinical service and some local and external clinical services. He is a member of the specialist advisory group for the National External Quality Assurance Service (NEQAS) for bone biomarkers. Neither of these roles are considered a conflict of interest. None of the other authors has any competing interests to declare.

Published February 2014
DOI: 10.3310/hta18110
This report should be referenced as follows:


*Health Technology Assessment* is indexed and abstracted in *Index Medicus/MEDLINE, Excerpta Medica/EMBASE, Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine.*
Criteria for inclusion in the Health Technology Assessment journal

Reports are published in Health Technology Assessment (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in Health Technology Assessment are termed ‘systematic’ when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. ‘Health technologies’ are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/15/01. The contractual start date was in July 2010. The draft report began editorial review in September 2012 and was accepted for publication in December 2012.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors’ report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen’s Printer and Controller of HMSO 2014. This work was produced by Burch et al. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).
Editor-in-Chief of Health Technology Assessment and NIHR Journals Library

Professor Tom Walley  Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein  Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May  Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key  Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck  Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke  Professor of Health Sciences, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly  Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson  Director of NETSCC, HTA, UK

Ms Tara Lamont  Scientific Advisor, NETSCC, UK

Professor Elaine McColl  Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire  Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads  Honorary Professor, Business School, Winchester University and Medical School, University of Warwick, UK

Professor Jane Norman  Professor of Maternal and Fetal Health, University of Edinburgh, UK

Professor John Powell  Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery  Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma  Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts  Professorial Research Associate, University College London, UK

Professor Helen Snooks  Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk
Abstract

Systematic review of the use of bone turnover markers for monitoring the response to osteoporosis treatment: the secondary prevention of fractures, and primary prevention of fractures in high-risk groups

Jane Burch,1 Stephen Rice,1 Huiqin Yang,1 Aileen Neilson,1 Lisa Stirk,1 Roger Francis,2 Paul Holloway,3 Peter Selby4 and Dawn Craig1*

1Centre for Reviews and Dissemination, York, UK
2Institute for Ageing and Health, Newcastle University, Newcastle upon Tyne, UK
3Imperial College Healthcare Trust, St Mary’s Hospital, London, UK
4Department of Medicine, Manchester Royal Infirmary, Manchester, UK

*Corresponding author

Background: There is currently no standard practice for the monitoring of patients receiving treatment for osteoporosis. Repeated dual-energy X-ray absorptiometry (DXA) is commonly used for monitoring treatment response, but it has its limitations. Bone turnover markers have advantages over DXA as they are non-invasive, relatively cheap and can detect changes in bone turnover rates earlier. However, they do have disadvantages, particularly high within- and between-patient variability. The ability of bone turnover markers to identify treatment non-responders and predict future fracture risk has yet to be established.

Objectives: We aimed to determine the clinical effectiveness, test accuracy, reliability, reproducibility and cost-effectiveness of bone turnover markers for monitoring the response to osteoporosis treatment.

Data sources: We searched 12 electronic databases (including MEDLINE, EMBASE, The Cochrane Library and trials registries) without language restrictions from inception to March 2012. We hand-searched three relevant journals for the 12 months prior to May 2012, and websites of five test manufacturers and the US Food and Drug Administration (FDA). Reference lists of included studies and relevant reviews were also searched.

Review methods: A systematic review of test accuracy, clinical utility, reliability and reproducibility, and cost-effectiveness of two formation and two resorption bone turnover markers, in patients being treated for osteoporosis with any of bisphosphonate [alendronsate (Fosamax®, MSD), risedronate (Actonel®, Warner Chilcott Company), zolendronate (Zometa®, Novartis)], raloxifene (Evista®, Eli Lilly and Company Ltd), strontium ranelate (Protelos®, Servier Laboratories Ltd), denosumab (Prolia®, Amgen Ltd) or teriparatide (Forsteo®, Eli Lilly and Company Ltd), was undertaken according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. Given the breadth of the review question, a range of study designs and outcome measures were eligible. The development of a decision model was planned to determine the cost-effectiveness of bone turnover markers for informing changes in patient management if clinical effectiveness could be established.

Results: Forty-two studies (70 publications) met the inclusion criteria; none evaluated cost-effectiveness. Only five were randomised controlled trials (RCTs); these assessed only the impact of bone marker monitoring on aspects of adherence. No RCTs evaluated the effectiveness of bone turnover marker
monitoring on treatment management. One trial suggested that feedback of a good response decreased non-persistence [hazard ratio (HR) 0.71, 95% confidence interval (CI) 0.53 to 0.95], and feedback of a poor response increased non-persistence (HR 2.22, 95% CI 1.27 to 3.89); it is not clear whether or not the trial recruited a population representative of that seen in clinical practice. Thirty-three studies reported results of some assessment of test accuracy, mostly correlations between changes in bone turnover and bone mineral density. Only four studies reported on intra- or interpatient reliability and reproducibility in treated patients. Overall, the results were inconsistent and inconclusive, owing to considerable clinical heterogeneity across the studies and the generally small sample sizes. As clinical effectiveness of bone turnover monitoring could not be established, a decision-analytic model was not developed.

**Conclusions:** There was insufficient evidence to inform the choice of which bone turnover marker to use in routine clinical practice to monitor osteoporosis treatment response. The research priority is to identify the most promising treatment–test combinations for evaluation in subsequent, methodologically sound, RCTs. In order to determine whether or not bone turnover marker monitoring improves treatment management decisions, and ultimately impacts on patient outcomes in terms of reduced incidence of fracture, RCTs are required. Given the large number of potential patient population–treatment–test combinations, the most promising combinations would initially need to be identified in order to ensure that any RCTs focus on evaluating those strategies. As a result, the research priority is to identify these promising combinations, by either conducting small variability studies or initiating a patient registry to collect standardised data.

**Funding:** The National Institute for Health Research Health Technology Assessment programme.
Contents

List of tables xi
List of figures xiii
List of abbreviations xv
Scientific summary xvii

Chapter 1 Background 1
Description of health problem 1
  Osteoporosis 1
  Bone turnover (remodelling) 1
  Diagnosis 1
Risk of fracture 1
Treatments for osteoporosis 2
Burden of the disease on the NHS 3
  National Institute for Health and Care Excellence guidance 4
Description of the technologies under assessment 4
  Bone turnover markers 4
  Variability in bone turnover markers 5
  Use of bone turnover markers 6
  Monitoring response to treatments for osteoporosis 6
Cost of the technologies under assessment 8
Summary 8

Chapter 2 Definition of decision problem 9
Decision problem 9
Overall aims and objectives of the assessment 9

Chapter 3 Assessment of the clinical effectiveness and cost-effectiveness evidence 11
Methods for reviewing the evidence 11
  Identification of studies 11
  Inclusion and exclusion criteria 12
  Data extraction strategy 14
  Critical appraisal strategy 14
  Methods of data synthesis 14
Results of the evaluation of clinical effectiveness 15
  Quantity and overall quality of research available 15
  Assessment of clinical effectiveness 24
  Assessment of test accuracy 26
  Assessment of test reliability and reproducibility 45
Results of the systematic review of cost-effectiveness 46
Discussion 46
  Key findings 47
  Strengths and limitations of the review 48
List of tables

TABLE 1 Summary of study characteristics 16

TABLE 2 Results of the RCTs reporting compliance as an outcome in patients who were and were not receiving feedback from bone turnover marker tests 25

TABLE 3 Score from the OPPS questionnaire as reported in Kung et al. (2009) 26

TABLE 4 Results from the studies reporting correlations between changes in bone turnover marker tests and either changes in DXA, vertebral strength index, or the incidence of fracture in patients being treated with bisphosphonates 27

TABLE 5 Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with bisphosphonates 30

TABLE 6 Results from studies using various methods to measure the predictive ability of bone turnover markers to predict changes in BMD or fracture risk in patients being treated with bisphosphonates 32

TABLE 7 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD, bone biopsy results, or the incidence of fracture in patients being treated with teriparatide 20µg/day 34

TABLE 8 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD in patients being treated with teriparatide 40µg/day 37

TABLE 9 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD or bone biopsy results in patients being treated with 20 or 40µg/day teriparatide, or where the dose was unclear 38

TABLE 10 Results from studies using various methods to measure the predictive ability of bone turnover markers to predict changes in BMD or fracture risk in patients being treated with teriparatide 40

TABLE 11 Results from studies reporting Pearson’s correlation coefficients for analyses associating changes in bone turnover markers and changes in BMD in patients being treated with raloxifene 41

TABLE 12 Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with raloxifene 42

TABLE 13 Results from studies reporting Spearman’s rank correlation coefficients for analyses associating changes in bone turnover markers and changes in BMD in patients being treated with strontium ranelate 43

TABLE 14 Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with strontium ranelate 44
TABLE 15  Results from studies using various methods to measure the predictive ability of bone turnover markers to predict changes in fracture risk in patients being treated with strontium ranelate 45

TABLE 16  Results from studies reporting Spearman’s rank correlations coefficients for analyses associating changes in bone turnover markers and changes in BMD in patients being treated with denosumab 45

TABLE 17  Signal to noise ratios reported for three bone turnover markers across studies with diverse populations and treatment regimens 46
List of figures

**FIGURE 1** Flow of studies through the review 15

**FIGURE 2** Summary of the quality of the non-randomised studies 23

**FIGURE 3** Current and feasible monitoring strategies for response to treatment 59
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>area under the curve</td>
</tr>
<tr>
<td>BALP</td>
<td>bone-specific alkaline phosphatase</td>
</tr>
<tr>
<td>BMD</td>
<td>bone mineral density</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>CG</td>
<td>clinical guideline</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CTX</td>
<td>carboxy-terminal telopeptide cross-linked type 1 collagen</td>
</tr>
<tr>
<td>CV</td>
<td>coefficient of variation</td>
</tr>
<tr>
<td>DXA</td>
<td>dual-energy X-ray absorptiometry</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>FDA</td>
<td>US Food and Drug Administration</td>
</tr>
<tr>
<td>GnRH</td>
<td>gonadotropin-releasing hormone</td>
</tr>
<tr>
<td>GP</td>
<td>general practitioner</td>
</tr>
<tr>
<td>HEED</td>
<td>Health Economic Evaluations Database</td>
</tr>
<tr>
<td>HR</td>
<td>hazard ratio</td>
</tr>
<tr>
<td>HRT</td>
<td>hormone replacement therapy</td>
</tr>
<tr>
<td>IRMA</td>
<td>immunoradiometric assay</td>
</tr>
<tr>
<td>ITT</td>
<td>intention to treat</td>
</tr>
<tr>
<td>MeSH</td>
<td>medical subject heading</td>
</tr>
<tr>
<td>MPR</td>
<td>medical possession ratio</td>
</tr>
<tr>
<td>NHS EED</td>
<td>NHS Economic Evaluation Database</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NTX</td>
<td>amino-terminal cross-linked type 1 collagen</td>
</tr>
<tr>
<td>ONJ</td>
<td>osteonecrosis of the jaw</td>
</tr>
<tr>
<td>OPPS</td>
<td>Osteoporosis Patient Perception Survey</td>
</tr>
<tr>
<td>OWH</td>
<td>Office for Women’s Health</td>
</tr>
<tr>
<td>P1CP</td>
<td>procollagen type 1 carboxy-terminal propeptide</td>
</tr>
<tr>
<td>P1NP</td>
<td>procollagen type 1 amino-terminal propeptide</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</td>
</tr>
<tr>
<td>PTH</td>
<td>parathyroid hormone</td>
</tr>
<tr>
<td>QALY</td>
<td>quality-adjusted life-year</td>
</tr>
<tr>
<td>QoL</td>
<td>quality of life</td>
</tr>
<tr>
<td>R²</td>
<td>regression coefficient (coefficient of determination)</td>
</tr>
<tr>
<td>RCT</td>
<td>randomised controlled trial</td>
</tr>
<tr>
<td>SAS</td>
<td>Supra-Regional Assay Service</td>
</tr>
<tr>
<td>sCTX</td>
<td>serum carboxy-terminal telopeptide cross-linked type 1 collagen</td>
</tr>
<tr>
<td>SERM</td>
<td>selective oestrogen receptor modulator</td>
</tr>
<tr>
<td>S/N</td>
<td>signal to noise ratio</td>
</tr>
<tr>
<td>sNTX</td>
<td>serum amino-terminal telopeptide cross-linked type 1 collagen</td>
</tr>
<tr>
<td>SPC</td>
<td>summary of product characteristics</td>
</tr>
<tr>
<td>TA</td>
<td>technology appraisal</td>
</tr>
<tr>
<td>T-score</td>
<td>number of standard deviations above/below mean for healthy 30-year-olds of same gender and ethnicity as the patient</td>
</tr>
<tr>
<td>uCTX</td>
<td>urinary carboxy-terminal telopeptide cross-linked type 1 collagen</td>
</tr>
<tr>
<td>uNTX</td>
<td>urinary amino-terminal telopeptide cross-linked type 1 collagen</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>Z-score</td>
<td>number of standard deviations above/below mean for patient’s age, gender and ethnicity</td>
</tr>
</tbody>
</table>
Scientific summary

Background

Osteoporosis is a progressive systemic skeletal disease characterised by low bone mass and micro-architectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture. Approximately 3 million people in the UK have osteoporosis, with about 20% of women aged 60–69 years being affected. There are approximately 230,000 osteoporotic fractures every year. Medical therapies available for osteoporosis include bisphosphonates, raloxifene, strontium ranelate, teriparatide and denosumab.

There is currently no standard practice for the monitoring of patients receiving treatment for osteoporosis. Repeated dual-energy X-ray absorptiometry (DXA) is a commonly used diagnostic test for monitoring treatment response but has its limitations, including the time needed prior to a repeated measure to detect changes in bone mineral density (BMD); limited access to the technology; cost (average £72 per scan); and evidence of the limited value in regular monitoring of BMD in patients on bisphosphonate therapy.

Bone turnover markers may offer an alternative monitoring strategy. They measure bone resorption or formation. Bone turnover markers have advantages over DXA for monitoring response to osteoporosis therapy; they are non-invasive, relatively cheap (commonly £20 to £25 per test), and have the ability to detect changes in bone turnover rates as early as 2 weeks for some therapies, and between 3 and 6 months for most. However, they do have disadvantages, most notably the variability across samples (both within and between patients). This leads to the need for a proportionately high percentage change in the rate of the bone turnover marker being measured in order to identify treatment responders. In addition, their ability to identify treatment non-responders and their use as independent predictors of future fracture risk has yet to be established.

Objectives

The primary aims of this assessment are to determine the clinical effectiveness, test accuracy, test reliability and reproducibility, and cost-effectiveness of monitoring regimens with at least one of four bone turnover markers, namely procollagen type 1 amino-terminal propeptide (P1NP), bone-specific alkaline phosphatase (BALP), carboxy-terminal telopeptide cross-linked type 1 collagen (CTX) and type 1 collagen amino-terminal telopeptide (NTX), in patients with osteoporosis being treated with any of bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide.

Methods

The review was conducted systematically following the general principles recommended in the Centre for Reviews and Dissemination’s guidance for undertaking reviews in health care and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement. Data were sought systematically from 12 electronic databases (including MEDLINE, EMBASE and The Cochrane Library) from inception up to March 2012. These were supplemented by searches of reference lists of included studies and relevant reviews, recent contents pages of relevant journals, and relevant websites. Inclusion was restricted to studies in adults (> 18 years of age) but not by date or language of publication.

To be included in the review, a study had to be either (1) a randomised controlled trial (RCT) comparing a monitoring regimen that included at least one bone turnover marker test with a monitoring regimen...
without bone turnover marker testing, or a different bone turnover marker, and reporting either change in patient management strategies and/or treatment adherence rates; (2) a study evaluating the impact of bone turnover marker test results on the decision-making process, that also reported the subsequent rate of fracture in the population; (3) a prospective study that compared the results of bone turnover marker tests with the results of bone biopsy or a composite reference standard of BMD and subsequent fracture outcome; (4) a prospective study that reported at least a $p$-value for the association between changes in bone turnover markers and BMD, biopsy, and/or the incidence of fractures from correlation or multivariate regression analyses; (5) a prospective study reporting inter- and/or intrapatient variability on bone turnover marker test results for patients receiving one of the treatments being evaluated; or (6) a cost-effectiveness analysis of bone turnover marker monitoring strategies. Non-effectiveness prospective studies had to recruit at least 20 patients with osteoporosis who were receiving one of the treatments of interest.

An economic model was to be developed only if sufficient evidence was found to establish the clinical effectiveness of bone turnover marker monitoring on treatment management.

**Results**

Forty-two studies (across 70 publications) met the inclusion criteria, all of which were included in the review of clinical effectiveness. Of the 42 studies, five were RCTs. Of the 37 non-randomised studies, 21 were cohorts derived from the treatment arms of RCTs, 15 were uncontrolled cohort studies and one was a controlled cohort study. All included studies were judged to be low quality. The high level of clinical heterogeneity across the studies precluded the use of standard meta-analytic techniques. A narrative synthesis was therefore employed.

**Clinical effectiveness**

Five RCTs and one post hoc analysis from a RCT assessed the effectiveness of feedback of bone turnover marker results on adherence, compliance and/or persistence. Five trials reporting on compliance showed little difference between the feedback and no feedback arms: high rates of baseline compliance mean that these are unlikely to be representative of clinical practice. Only one trial reported on persistence. Notably, feedback of a good urinary NTX (uNTX) response (> 30% reduction) was associated with a decreased rate of discontinuation [hazard ratio (HR) 0.71, 95% confidence interval (CI) 0.53 to 0.95]. In contrast, feedback of a poor uNTX response was associated with an increased rate of discontinuation (HR 2.22, 95% CI 1.27 to 3.89). Two RCTs reported on the quality of life (QoL) using the osteoporosis-specific questionnaire; these variably reported small improvements for patients receiving feedback in the overall, feeling informed, satisfaction and confidence scores. No studies were identified for the evaluation of the effectiveness of bone turnover marker monitoring on treatment management.

**Test accuracy**

Thirty-three studies reported results of some assessment of test accuracy, 23 reported only the results of correlation analyses, four only the results of multiple regression analyses, and four reported both. Five studies reported predictive accuracy using alternative analytical methods; three also reported results from correlation and/or multiple regression analyses. Therefore, most of the data identified for the review of test accuracy were results from correlation analyses; the majority of these evaluated associations between changes in bone turnover markers with changes in BMD. Although there were a number of statistically significant associations between these two measures across the different treatments, the vast majority had small effect sizes and were considered weak ($r < 0.50$). The studies that used regression analyses to adjust for confounding factors gave some indication that changes in bone turnover markers may be significantly associated with subsequent changes in BMD. However, there were too few of these studies to draw any firm conclusions. Studies assessing the association between changes in bone turnover markers with either biopsy results or fracture outcomes were uncommon. Two studies used biopsy and seven used fracture, and these gave some indication that changes in bone turnover markers may be significantly associated with changes in fracture risk; however, again, there were too few studies to draw any firm conclusions.
Overall, the results from the studies utilising correlation and regression analyses were inconsistent and inconclusive. This may be due to the considerable clinical heterogeneity across the included studies in terms of the definitions used to identify those with osteoporosis, patient populations recruited, the treatment regimens administered, and the type and timing of the tests being evaluated. Most of the included studies had small sample sizes, resulting in low statistical power to detect significant associations.

**Test reliability and reproducibility**

Four studies reported signal to noise (S/N) ratios for a bone turnover marker in patients being treated with etidronate, teriparatide or raloxifene. Within-study comparisons showed that serum P1NP (sP1NP) had a higher S/N ratio than serum CTX (sCTX) at 25 weeks, and a higher S/N ratio than serum BALP (sBALP) at 6 months.

**Cost-effectiveness**

No studies met the inclusion criteria for the systematic review of the cost-effectiveness of bone turnover marker monitoring strategies.

**Economic model**

Given that the review could not establish the clinical effectiveness of bone turnover marker monitoring strategies, a decision-analytic model could not be produced and, consequently, an expected value of perfect information could not be undertaken to assess the value of future research.

To assist future developers of any decision-analytic model in investigating the cost-effectiveness of bone turnover marker monitoring strategies, we undertook a scoping review of current modelling methods in related decision problems. We also discussed the gaps in the current evidence base that would be essential to address before any such cost-effectiveness analysis of bone marker monitoring regimens could be undertaken.

Of the modelling strategies identified, 12 modelled measures of adherence and one modelled treatment change. Ten of the models incorporated compliance as a binary variable, using a variety of cut-off points for what constituted compliance. Eleven models incorporated persistence, modelled as the percentage of patients initiating and subsequently discontinuing treatment at different time points. Only six studies modelled compliance, non-compliance and persistence separately, incorporating the different aspects of adherence. Some models included an estimate of primary non-adherence. The one model that incorporated treatment change allowed for switching to a second-line treatment if results of a bone turnover marker test during follow-up led to the conclusion that compliance or response to treatment was inadequate.

The key part of any future cost-effectiveness analysis of bone turnover marker tests for monitoring response to treatment for osteoporosis is accounting for test accuracy, the prognostic outcomes for true-positive, false-positive, true-negative and false-negative test results, and the effect of feeding back the results of bone turnover marker tests on patient adherence to treatment. These data were either absent completely, insufficient given the different tests and treatments, or applicable to populations with unrealistic adherence rates for clinical practice.

**Discussion**

The systematic review of clinical effectiveness found no evidence evaluating the impact of treatment monitoring regimens that included a relevant bone turnover marker on treatment management decisions. The review identified limited data assessing the effect of bone turnover marker feedback on patient compliance, persistence and/or adherence to treatment, the results of which suggested that the positive feedback results encouraged patient persistence.
Most of the data relating to test accuracy were in the form of correlations between changes in bone turnover markers (usually between 1 month and 6 months of starting treatment) and subsequent changes in BMD (usually between 1 year and 3 years after the start of treatment). Treatment-induced changes in BMD account for a limited proportion of the observed reduction in fracture risk and, therefore, BMD is a poor surrogate for fracture risk; using BMD as a surrogate for the evaluation of the predictive accuracy of bone turnover markers to identify patients on treatment who remain at risk of fracture is inappropriate. In addition, results of correlation analyses are influenced by sample size: the greater the sample size, the more likely a correlation will be statistically significant from zero. Although there were a number of statistically significant correlations, these on the whole suggested weak correlations. These data, and the data from studies conducting multiple regression analyses, were further limited by the considerable between-study clinical heterogeneity in terms of the definitions of osteoporosis, patient populations, treatment regimens and the type and timing of tests being evaluated.

In terms of the evaluation of test reliability and reproducibility, some evidence was available that suggested sP1NP may have a greater S/N ratio than sBALP and sCTX at a short-term follow-up, but the data on this outcome were sparse and longer-term follow-up data absent.

The systematic review of cost-effectiveness identified no studies evaluating different treatment monitoring strategies, where BALP, P1NP, CTX or NTX was incorporated as part of one of the strategies, and there was insufficient evidence from the clinical review to develop a de novo decision-analytic model.

Overall, the evidence required to address the decision problem was lacking. The evidence that was available was heterogeneous and of poor quality. Consequently, it was impossible to draw any conclusion as to whether or not bone turnover markers were able to identify non-responders or predict fracture risk independently of BMD in patients receiving osteoporosis treatment. There are a number of uncertainties that remain in need of clarifying; these include:

- the ability of changes in bone turnover markers to identify treatment non-responders
- the ability of changes in bone turnover markers to impact on compliance, persistence and adherence to each of the treatments being evaluated
- the accuracy of changes in bone turnover markers to predict future fracture risk
- the ability of bone turnover markers to inform treatment change
- the most appropriate timing of the conduct of bone turnover marker testing; this may vary depending upon the treatment–test combination
- which bone turnover marker is superior in terms of its ability to identify treatment non-responder and predict fracture risks for monitoring specific osteoporosis treatments
- the reliability and reproducibility of bone turnover marker tests in patients receiving treatment for osteoporosis
- the most cost-effective monitoring regimen for patients being treated with bisphosphonates, raloxifene, strontium ranelate, teriparatide or denosumab.

**Conclusions**

**Implications for service provision**

The lack of evidence of clinical effectiveness and the heterogeneity and poor quality of the available evidence on the accuracy, reliability and reproducibility of bone turnover markers for monitoring response to osteoporosis treatment precluded the possibility of making any recommendations on the choice of bone turnover marker being used in routine clinical practice for its superiority to monitor osteoporosis treatment response. In addition, the evidence to support the use of bone turnover marker feedback results to improve patient adherence to osteoporosis treatment was not convincing.
**Suggested research priorities**

In order to determine whether or not bone turnover marker monitoring improves treatment management decisions and ultimately impacts on patient outcomes in terms of reduced incidence of fracture, RCTs are required. The predictive accuracy of bone turnover markers for future fracture outcomes in patients receiving osteoporosis treatment could be investigated using prospective, long-term observational studies with large sample sizes. However, in view of the large number of potential patient population–treatment–test combinations, the most promising combinations would need to be identified in order to ensure the more costly and time-consuming studies, such as RCTs, focus on evaluating those strategies. Therefore, we consider the research priority to identify these promising treatment–test combinations. This can be achieved by either conducting small variability studies or initiating a patient registry to collect standardised data. The former would be quicker, easier and less costly, but the quality of the data would be poorer. Further, prior to establishing the latter it is likely that a more widespread use of bone turnover markers in clinical practice would be required. Once the most promising treatment–test combinations have been identified, well-designed RCTs can be conducted to evaluate the effectiveness of those monitoring regimens; this would include measuring outcomes such as the proportion of non-responders, adherence rates, treatment management decisions and fracture outcome. Data from these RCTs along with other sources can then be included in a decision-analytic model in order to investigate cost-effectiveness.

**Funding**

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.
Chapter 1  Background

Description of health problem

Osteoporosis
Osteoporosis is a progressive systemic skeletal disease characterised by low bone mass and micro-architectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture.1,2

Bone turnover (remodelling)
Bone turnover is the process of resorption followed by replacement by new bone with little change in shape, and it occurs throughout a person’s life. Osteoclasts break down bone (bone resorption), releasing the minerals, resulting in a transfer of calcium from bone fluid to the blood. The osteoclast attaches to the osteon (layers of compact bone tissue surrounding a central canal), and secretes collagenase and other enzymes. Calcium, magnesium, phosphate and products of collagen are released into the extracellular fluid as the osteoclasts tunnel into the mineralised bone. Osteoblasts are mature bone cells responsible for bone formation and ossification. They produce the organic portion of the matrix of bone tissue, osteoid, which is composed mainly of type I collagen, and are responsible for mineralisation of the osteoid matrix.

Ossification fixes circulating calcium in its mineral form, removing it from the bloodstream. Repeated stress, such as weight-bearing exercise or bone healing, results in the bone thickening at the points of high stress.

Remodelling in adults repairs micro-damage to bone and plays a role in the regulation of calcium homeostasis. An imbalance in the bone remodelling processes in adults is thought to impact on bone strength as a result of reductions in bone volume and mineralisation, loss of trabeculae, deterioration of trabecular connectivity, and the formation of resorption cavities and trabecular perforations.3,4 Therefore, an increase in bone turnover where resorption exceeds formation is not only inversely correlated with bone mineral density (BMD), but may also alter bone architecture and porosity, increasing the risk of fracture beyond that due to reduced BMD, and can therefore be an independent predictor of fracture risk.3–6

Diagnosis
Osteoporosis causes no symptoms until a bone is broken. As osteoporosis is associated with low bone density, bone density scanning [using dual-energy X-ray absorptiometry (DXA)] has become the most commonly used diagnostic technique.2 There are accepted diagnostic criteria based on DXA: osteopenia (low bone mass) is present when the BMD is between 1 and 2.5 standard deviations below the mean value for young adults (BMD T-score of –1 to –2.5); osteoporosis is diagnosed when BMD is < 2.5 standard deviations below young adults’ (BMD T-score of < –2.5).7

Risk of fracture
A reduction in BMD results in the thinning of the trabeculae and an increase in the fragility of the bones.8 Therefore, people diagnosed with osteoporosis have an increased risk of suffering low trauma (fragility) fractures. When BMD is measured by DXA, a reduction of 1 standard deviation in BMD is reportedly associated with a 50–150% increase in the risk of osteoporotic fracture.9 Increasing age is one of the major risk factors for osteoporosis; after 35 years of age bone loss increases gradually as part of the natural ageing process.2 By 75 years of age, approximately half of the population will have osteoporosis. In addition, there is an increased risk of falling which increases the risk of fracture; one in two women and one in five men over the age of 50 in the UK will fracture a bone, mainly as a result of skeletal fragility.2,10 The most common fractures in people with osteoporosis are of the wrists, hips and spinal bones; these are most common in older people, but younger people can sometimes be affected.8,11
According to recent National Institute for Health and Care Excellence (NICE) guidance [clinical guideline (CG) 146], assessment of the risk of fragility fractures should be considered in:

- all women aged 65 years and over and in all men aged 75 years and over
- in women aged under 65 years and in men aged under 75 years in the presence of risk factors, for example:
  - previous fragility fracture
  - current use or frequent recent use of oral or systemic glucocorticoids
  - history of falls
  - family history of hip fracture
  - other causes of secondary osteoporosis
  - low body mass index (BMI) (< 18.5 kg/m²)
  - smoking
  - alcohol intake of more than 14 units per week for women and more than 21 units per week for men.

An assessment tool for assessing fracture risk, FRAX®, has been developed by the World Health Organization (WHO). The factors taken into account are age, gender, weight, height, previous fracture, parental history of hip fracture, smoking status, the use of oral glucocorticoid steroids, a diagnosis of rheumatoid arthritis, the presence of a disorder strongly associated with osteoporosis and alcohol consumption, with or without BMD as determined using DXA.

**Treatments for osteoporosis**

Diet and exercise can be modified to improve a person’s fracture risk. Exercises considered best for people with osteoporosis are those that (1) are thought to have an effect on density and strength, such as weight-bearing exercises that cause force on the bones like jogging, stair climbing, walking briskly and resistance exercises, and (2) can reduce the risk of falling, such as balance training (e.g. tai chi), leg strengthening and flexibility training (e.g. yoga). Exercises that people with osteoporosis are advised to avoid are those that might increase the risk of falling, those that involve twisting the spine or bending from the waist, high-impact activities such as high-intensity aerobics or jumping and the use of excessive weight during resistance exercise. A diet containing foods rich in calcium and vitamin D (vitamin D is required for the absorption of calcium) or the use of calcium and vitamin D supplements can also improve bone strength.

The most common medical therapies for osteoporosis are bisphosphonate drugs. Bisphosphonates inhibit the activity of mature osteoclasts and reduce the rate of resorption. The most commonly prescribed bisphosphonate is generic alendronate; other bisphosphonates include etidronate, risedronate (now available in generic form), ibandronate, and zoledronate. The recommended dose of alendronate is one 70-mg tablet per week, rather than 10 mg daily as originally prescribed, to reduce the incidence of gastrointestinal adverse effects and increase adherence. A strict technique must be adhered to when taking oral bisphosphonates to ensure satisfactory absorption. They must be taken on an empty stomach first thing in the morning, while remaining upright to prevent reflux, at least 30 minutes before the first food, drink or other medication of the day. The tablet should be taken with plain water only; other drinks (including mineral water), food and some medicines are likely to reduce the absorption of bisphosphonates. Intravenously administered bisphosphonates are available; the recommended doses are 3 mg 3-monthly of ibandronate, or 5 mg annually of zoledronate. Pamidronate is not licensed for the treatment of osteoporosis but has been widely used off-licence at a dose of 30 mg quarterly.

Other medical therapies available include:

- raloxifene (Evista®, Eli Lilly and Company Ltd): a selective oestrogen receptor modulator (SERM), which is a synthetic hormone that copies the effects of oestrogen on the bones
• **strontium ranelate** (Protelos®, Servier Laboratories Ltd): a strontium(II) salt of ranelic acid, which is a dual-action bone agent that stimulates new bone growth and reduces bone loss

• **teriparatide** (Forsteo®, Eli Lilly and Company Ltd): a recombinant form of parathyroid hormone (PTH 1–34) that helps regulate calcium levels and the activity of cells involved in bone formation

• **denosumab** (Prolia®, Amgen Ltd): a monoclonal antibody that targets the RANK ligand

• **hormone replacement therapy (HRT)**: a mix of hormones (oestrogens, progesterone or progestins, and sometimes testosterone) prescribed to post-menopausal women (natural or surgically induced) to reduce the symptoms caused by reduced circulating oestrogen and progesterone. The risk of development and progression of osteoporosis can therefore be reduced by the maintenance of oestrogen levels.

### Burden of the disease on the NHS

Approximately 3 million people in the UK have osteoporosis, with about 20% of women aged 60–69 affected. There are thought to be about 230,000 osteoporotic fractures every year, with broken wrists, hips and spinal bones being the most common. Of the 60,000 people who suffer osteoporotic hip fractures each year, 15–20% are likely to die within a year from causes related to the fracture.²

As stated in *Diagnosis*, above, there are a range of treatments available for osteoporosis, and the costs of these vary (pamidronate has not been costed as it is not licensed for use in osteoporosis):¹⁶

- **Generic sodium alendronate**: a 28-tablet pack of 10-mg tablets is £1.44 (approximately £19 annually); a four-tablet pack of 70 mg for once-weekly administration is £1.10 (approximately £14 annually). Fosamax® (MSD) costs £23.12 for 28 10-mg tablets and £22.80 for four 70-mg once-weekly tablets.
- **Generic sodium risedronate**: a 28-tablet pack of 5-mg tablets is £17.99 (approximately £220 annually); a four-tablet pack of 35 mg for once-weekly administration is £19.12 (approximately £230 annually).
- **Zoledronate**: Zometa® (Novartis) costs £174.17 for 4 mg in 5 ml, and Aclasta® (Novartis) costs £253.38 for 5 mg in 100 ml – 5 mg administered once annually.
- **Strontium ranelate** (Protelos®, Servier) costs £25.60 for 28 sachets each containing 2 g of granules daily (approximately £330 annually).
- **Denosumab** (Prolia®, Amgen Ltd) costs £183.00 for 60 mg/ml in a 1-ml prefilled syringe – 60 mg administered 6-monthly (£366 annually).
- **Raloxifene** (Evista®, Daiichi Sankyo) costs £17.06 for 28, and £59.59 for 84, 60-mg tablets – 60 mg daily (approximately £220 annually).
- **Teriparatide** (Forteo®, Eli Lilly and Company Ltd) costs £271.88 for 250 µg/ml in a 3-ml pre-filled pen – 20 µg self-administered daily (approximately £3540 annually).

According to Hospital Episode Statistics (HES), in 2005–6 in England there were 5759 consultations and 4034 admissions (2368 emergency) for osteoporosis with a pathological fracture, and a further 8725 consultations and 8313 admissions (716 emergency) without pathological fracture.¹⁷ For surgical interventions for fractures of the spine and hip (not only those associated with osteoporosis), there were 809 consultations and 667 admissions (353 emergency) for fixations of spinal fractures (approximately 26% in patients 60 years and older), and 46,812 consultations and 46,191 admissions (1611 emergency) for primary total prosthetic replacement of hip joint [depending on method used, approximately 50% (not using cement) to 85% (using cement) 60 years and older].¹⁷ Given the discrepancies in the numbers of hip replacements in the elderly and consultations of osteoporotic fractures, the incidence/consultation rate for osteoporosis may be underestimated. A recent report published by the Royal College of Physicians stated that only 32% (1933 out of 6083) of non-hip fracture and 67% (2324 out of 3484) of hip fracture patients had a clinical assessment for osteoporosis/fracture risk.¹⁸ Osteoporosis reportedly costs the NHS and government £2.38 per year (£6M per day).²
National Institute for Health and Care Excellence guidance

NICE has produced a number of technology appraisals (TAs) and CGs that have some relevance to this area. Three relevant TAs have been published: TA160 (Osteoporosis – primary prevention; postmenopausal women),19 TA161 (Osteoporosis – secondary prevention including strontium ranelate; postmenopausal women)20 and TA204 (Osteoporotic fractures – denosumab).21

For the primary prevention of fractures, alendronate is recommended as the first-line treatment for most women at risk of fractures. Risedronate, etidronate and strontium ranelate are alternative treatments for post-menopausal women who cannot adhere to the required alendronate regimen, or those women with pre-specified combinations of T-score, age and number of independent clinical risk factors; strontium ranelate is not recommended as a first-line treatment for osteoporosis. Raloxifene is not a recommended treatment for the primary prevention of osteoporotic fragility fractures.19 The recommendations for the secondary prevention of fractures are similar to those for primary prevention. The two differences are that (1) strontium ranelate can be used as a first-line treatment and (2) raloxifene is recommended as an alternative treatment for post-menopausal women who cannot adhere to alendronate, or in women with pre-specified combinations of T-score, age and number of independent clinical risk factors.20 Denosumab has now also been added to the list of alternative second-line treatments for the primary or secondary prevention of fractures.21

There are also four potentially relevant CGs available that deal with the management of independent risk factors for fracture: CG146 (Osteoporosis fragility fracture),12 CG21 (Falls: the assessment and prevention of falls in older people),22 CG59 (Osteoarthritis: the care and management of osteoarthritis in adults)23 and CG79 (Rheumatoid arthritis: the management of rheumatoid arthritis in adults).24

This review will focus on patients being treated for osteoporosis with any of bisphosphonate, raloxifene, strontium ranelate, teriparatide or denosumab.

Description of the technologies under assessment

Bone turnover markers

Biochemical markers of bone turnover are used to monitor treatment response and may prove to be more useful than serial BMD measurements as they are non-invasive, relatively cheap compared with DXA, and there is an increased availability of auto-analysers in clinical chemistry laboratories.

Formation markers (detects products from the action of osteoblasts)

Bone-specific alkaline phosphatase (BALP): serum alkaline phosphatase has several dimeric isoforms that originate from a range of tissues (liver, bone, intestine, spleen, kidney and placenta), with approximately 40–50% of the total alkaline phosphatase activity arising from the bone as a result of osteoblast activity.25 The bone-specific isoform can be detected with immunoassays using monoclonal antibodies.26,27 There are two main types of assay to measure BALP: enzyme-linked immunosorbent assay (ELISA; measures BALP enzyme activity) and immunoradiometric assay (IRMA; measures BALP in protein mass units).28 The least significant change between a sample taken at baseline to 3 months after commencement of treatment has been reported as 30%.27 It has been suggested that BALP testing should occur at baseline before starting osteoporosis therapy and again at 3 to 6 months after commencement of therapy.29

Procollagen type 1 amino-terminal propeptide (P1NP): anti-P1NP antibodies are used to detect the trimeric structure of P1NP by ELISA or radioimmunoassay. It has been claimed that P1NP is a more sensitive marker of bone formation rate than other available formation markers, and therefore is particularly useful for monitoring bone formation therapies and antiresorptive therapies.26,29 As with BALP, it is recommended that the test be performed at baseline before starting osteoporosis therapy and again 3–6 months later.29
Osteocalcin (or bone gla protein): a small protein, detected using ELISA or radioimmunoassay that is rapidly degraded in the serum so that intact and fragmented segments from osteoblast activity coexist in the serum. Advantages of osteocalcin have been reported as being its tissue specificity, wide availability, and relatively low within-person variation; however, heterogeneity of the fragments in the serum is thought to limit its use.\(^{26}\) Osteocalcin is a marker of corticosteroid effects on osteoblasts and is decreased in patients receiving acute high-dose steroids, a risk factor for osteoporosis;\(^{27}\) osteocalcin may also be affected by use of warfarin.\(^{29}\) It is recommended that the test be performed at baseline before starting osteoporosis therapy and again 3–6 months later.\(^{29}\)

Procollagen type 1 carboxy-terminal propeptide (P1CP): the carboxy-terminal propeptide cleaved during the assembly of collagen fibres, and detected using ELISA or radioimmunoassay.\(^{30}\)

Resorption markers (detects products from the action of osteoclasts)

Carboxy-terminal telopeptide cross-linked type 1 collagen (CTX): peptide fragments from the carboxy-terminal end of type 1 collagen produced during osteoclastic resorption and detected in the urine or serum using ELISA.\(^{29}\)

Type I collagen amino-terminal telopeptide (NTX): peptide fragments from the amino terminal end of type 1 collagen produced during osteoclastic resorption and detected in the urine or serum with competitive inhibition ELISA or a chemiluminescence assay.\(^{27,29}\) The least significant change between samples taken at 3-month intervals is 50%. Suppression of NTX by more than 50% from baseline has been reported as being expected as early as 3 months after commencement of bisphosphonate therapy, but routine follow-up may be left to 6 months post therapy.\(^{27}\) It has been recommended that the test be performed at baseline before starting osteoporosis therapy and again 3 to 6 months later.\(^{29}\)

Urine deoxypyridinoline: derived only from bone matrix degradation, released from type I collagen. Excretion of deoxypyridinoline expressed as ratio to creatinine excretion. Urine deoxypyridinoline is detected by high-performance liquid chromatography or competitive ELISA.\(^{27}\) Increases of between two and three times the upper limits of normal have been reported in people with osteoporosis, primary hyperparathyroidism, osteomalacia, thyrotoxicosis and several inflammatory conditions, though the biggest increases (four or more times upper limit of normal) are seen in immobilisation, Paget’s disease of bone and metastatic cancer.\(^{27}\) A decrease in the pretreatment value of > 30% has been considered indicative of a good response in osteoporosis.\(^{27}\)

The Supra-Regional Assay Service (SAS) is a UK-based service for the analysis and clinical interpretation of a wide range of specialised diagnostics tests; those offering BALP, uNTX, serum osteocalcin and urine deoxypyridinoline are listed on the SAS website.\(^{27}\)

Variability in bone turnover markers

Several factors can impact on the bone turnover marker levels, causing variability across samples, which can reduce repeatability and comparability, both within patients and between patients. These include specimen collection and storage;\(^{25,31–35}\) differences between analytical methods used;\(^{32,34}\) temporal variations (diurnal, menstrual, seasonal);\(^{25,31–35}\) diet and fasting;\(^{36}\) patient characteristics (age, gender or ethnicity);\(^{25,31,33,35}\) concomitant medication other than osteoporosis medications [HRT, anabolic agents, glucocorticoids, anticonvulsants, gonadotropin-releasing hormone (GnRH) antagonists or oral contraception];\(^{25,31}\) and comorbid conditions (renal impairment, liver disease, diabetes, thyroid disease, osteomalacia, systemic inflammatory diseases, degenerative joint disease, conditions causing immobility, or eating disorders).\(^{25,31,33,35}\)

Intrapatien variability for serum markers is lower than for urinary markers.\(^{34}\) Some tests are more accurate when monitoring the response to specific treatments (e.g. CTX with bisphosphonates). Some tests have the advantage of not requiring the patient to fast prior to sampling (e.g. P1NP), or are less affected by diurnal variations (P1NP and BALP), and/or have lower overall intraindividual variability (BALP) than other
bone turnover markers. Each of these tests also has disadvantages: CTX has a large circadian rhythm, and therefore repeat sampling must be done at the same time of day, fasting is required prior to sampling, and the marker requires freezing soon after sampling as it can be unstable; BALP is affected by cross-reactivity with the liver form of alkaline phosphatase, limiting its use in patients with liver disease; and P1NP has a higher cost compared with other bone turnover markers. Given the advantages that CTX, P1NP and BALP offer, and the availability of NTX, these are the bone turnover markers that will be investigated in the current review.

**Use of bone turnover markers**
The use of bone turnover markers varies greatly across the UK, in terms of both the test used and the frequency of its measurement. Several factors will need to be considered when choosing the bone turnover marker to be used, not least the availability of the assay methods. Bone turnover markers have a number of potential uses, including:

1. predicting bone loss
2. identifying people at risk of primary or secondary osteoporosis and fracture
3. predicting treatment response prior to commencement
4. monitoring the response to osteoporosis treatment; identifying non-responders, which will include those not adhering with osteoporosis treatment (including patients not taking the medication or not following the instructions for administration)
5. identifying oversuppression of bone turnover in patient on long-term osteoporosis therapy
6. monitoring of people who have been on long-term treatment, or shown signs of oversuppression, and are taking a ‘treatment holiday’.

The main focus of this systematic review will be role 4: monitoring the response and non-response to osteoporosis therapy (and change in fracture risk).

**Monitoring response to treatments for osteoporosis**
There is currently no standard practice for the monitoring of patients receiving treatment for osteoporosis. The options include the use of repeated DXA, repeated measures of bone turnover markers, clinical review, or a combination of these. The use of DXA to monitor the response to osteoporosis treatment has limitations. Firstly, detectable changes in bone density due to treatment can take up to 2 years to become apparent, therefore, the identification of non-responders to treatment is delayed. Secondly, there is limited access to the technology and the test is relatively expensive (average £72 per scan). Thirdly, there is evidence that there is limited value in the regular monitoring of BMD in patients on bisphosphonate therapy.

As stated earlier, the relationship between bone turnover and bone density and architecture means that the rate of bone turnover may be an independent predictor of fracture risk; this can be measured using one or more of the bone turnover markers listed above. However, it is still unclear whether or not changes in bone turnover detected by bone turnover markers are reliable surrogate measures for improved bone density and architecture, and consequently accurate predictors of future fracture risk. Two studies have suggested that bone turnover markers can have independent predictive value in assessment of fracture risk. If biochemical markers of bone turnover are reliable indicators of future fracture risk, their use may prove advantageous compared with serial BMD measurements, as not only are they non-invasive, relatively cheap compared with DXA, and the availability of auto-analysers in clinical chemistry laboratories is increasing, but a response to treatment can be detected much earlier than with DXA.

Changes in bone turnover rates have been detected in post-menopausal women within as early as 2 weeks after starting HRT, although the peak accuracy of changes in bone turnover markers to predict fracture risk in response to osteoporosis treatment may be later than this, between 3 and 12 months after initiating treatment, depending on the treatment and bone turnover marker used. The ability to identify non-responders early within the treatment can be beneficial for patients by allowing early changes in management strategy if deemed necessary. The definition of treatment success varies depending upon...
the baseline risk of the patient being treated; in some patients a reduction in bone turnover would be considered a treatment success, but in others success may be a stabilisation of bone turnover. For all patients a continued increase in bone turnover rates would be considered a treatment failure. The definitions used throughout this project will reflect clinical practice and be based upon evidence for least clinical significant change.

There is a complex association between changes in bone turnover and fracture risk that is influenced by the treatment–bone turnover marker combination; the observed change in bone turnover markers will depend upon the treatment being administered. In studies of raloxifene, risedronate, alendronate and zoledronic acid, bone turnover markers have been reported as explaining between 28% and 77% of fracture risk reduction.47

Bisphosphonates are antiresorptive therapies, and therefore they reduce the rate of bone resorption. Bone resorption is closely coupled to bone formation; consequently, there is usually a subsequent reduction in the rate of bone formation. This results in a transient uncoupling of bone turnover, which leads to a small increase in BMD. This increase in BMD may account in part for the decrease in fracture risk, but the reduction in bone turnover may independently improve bone strength by improving bone architecture and porosity.48 Both raloxifene and denosumab reduce bone resorption, and therefore act as antiresorptive therapies; decreases in both bone resorption markers and subsequently bone formation markers should be observed in treatment responders as with bisphosphonates.

Teriparatide causes a small, transient increase in serum calcium, mainly due to the stimulation of tubular reabsorption of calcium from the proximal kidney tubules and increased calcium absorption from the bowel, but in a small part by increasing bone resorption (hence chronically elevated PTH can deplete bone). However, intermittent administration of PTH (i.e. daily injections of teriparatide) activates osteoblasts more than osteoclasts, stimulating new bone formation and increasing BMD. Therefore, a positive response in bone formation markers would be expected in treatment responders, with a subsequent increase in bone resorption markers due to the coupling of the processes, the opposite response to that seen with antiresorptive therapies.

Strontium ranelate increases new bone formation as well as reducing bone resorption and is classed as a dual-action bone agent. These effects are more modest than those seen with anabolic and antiresorptive treatments, with smaller positive changes in bone formation and negative changes in bone resorption markers, respectively. However, strontium ranelate appears to lead to persistent uncoupling of bone turnover.

The interpretation of changes in bone turnover markers is also influenced by the type of sample used: serum or urine. The intraindividual variability is greater for urine markers, giving serum markers a better signal to noise (S/N) ratio;34 the percentage change in a urinary biomarker needed to indicate a treatment response (least significant change) is greater than that required for a serum biomarker.

Treatment non-response
Treatment non-response could have a number of causes, including non-compliance; non-persistence; an underlying, untreated cause of the osteoporosis; an inability to absorb the drug; and/or test error. The most common reasons are thought to be non-compliance, non-persistence, or both (non-adherence).

Adherence to osteoporosis treatment is known to be poor, particularly to bisphosphonates, which are often associated with gastrointestinal upset and sometimes oesophagitis.49 According to the summary of product characteristics (SPC), gastrointestinal upset with alendronate is common (occurring in 1–10% of patients) and oesophagitis is rare (0.01–0.1% of patients).15 The incidence of gastrointestinal side effects associated with osteoporosis treatments is thought to be higher than that specified in the SPC; NICE guidance states that up to one-third of post-menopausal women may experience some type of gastrointestinal upset.50,51 The occurrence of more severe oesophageal complications reported in post-marketing surveillance has been
put down to taking alendronate with little or no water, lying down during or shortly after taking the tablet, continuing to take alendronate after the onset of symptoms, or pre-existing oesophageal disorders. Patients are now given strict instructions on the technique for taking bisphosphonate drugs, as described previously. Adverse events have been reported in nearly 50% of patients; however, a 2006 Cochrane review showed no significant difference in gastrointestinal adverse events between bisphosphonates and placebo. In addition to the potential for adverse events, bisphosphonates are difficult to absorb. Patients have to adhere to strict instructions on how to take oral preparations; if these are not followed, the effectiveness of the drug is likely to be reduced and gastrointestinal side effects are more likely to be experienced.

Bone turnover markers can identify treatment non-responders, and therefore they may be a useful method for monitoring non-adherence with treatment, as this is a major reason for non-response. Adherence to treatment can be improved with the introduction of treatment regimens that require less frequent administration of the medication, and the availability of intravenously administered bisphosphonates. The move to the use of intravenously administered treatment based on the results of the bone turnover markers could have cost implications; anaphylaxis could occur and, if experienced, it may require hospitalisation. Monitoring adherence through the use of bone turnover markers is not a main focus of the systematic review; however, where this information is reported it will be extracted and summarised.

Cost of the technologies under assessment

In England, in 2010–11, DXA cost, on average, £72 per scan (range £45 to £85: Health Resource Group code RA15Z). In comparison, a bone turnover marker assay can cost approximately £20 to £25; this includes administration and clinical interpretation costs as well as the cost of the reagents. P1NP had been reported as costing between £25 and £83 in 2007.

Summary

Bone turnover markers may be useful in monitoring the response of bone turnover to treatment regimens in patients with osteoporosis, and hence to identify patients who are non-responders. This in turn will allow changes in management or treatment strategies to be implemented in a timely manner to ensure maximum benefit to the patient. An evidence synthesis using systematic review methodology will be used to investigate potential uses of bone turnover markers and a decision-analytic model will be developed, if sufficient evidence is found, to establish clinical effectiveness.
Chapter 2 Definition of decision problem

Decision problem

In relation to the use of bone turnover markers for the monitoring of patients receiving osteoporosis treatments, the decision problem in clinical practice is: ‘What is the clinical and cost-effectiveness of monitoring regimens that include at least one bone turnover marker, over and above monitoring regimens where a bone turnover marker is not used, and which, if any, bone turnover marker should be introduced into routine practice for the monitoring of response to osteoporosis treatments?’

Overall aims and objectives of the assessment

The primary aims of the systematic review are to determine the clinical effectiveness, test accuracy, test reliability, test reproducibility and cost-effectiveness of bone turnover markers in people being treated with any of bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for osteoporosis.

The review of the clinical evidence will focus on three key clinical areas:

- Clinical effectiveness: how does bone turnover marker monitoring impact on the decision-making process and patient outcomes?
- Test accuracy: how well do changes in the level of bone turnover markers associate with changes in bone density, architecture and incidence of fracture?
- Test reliability and reproducibility: how much do the results of tests vary within and between patients?

If possible (i.e. if clinical effectiveness can be established) a decision model will be developed to determine the cost-effectiveness of bone turnover markers for monitoring treatment response and making changes in patient management, addressing the question: ‘Which monitoring regimen is the most cost-effective for informing treatment decisions?’ The treatments considered in the model will be those considered in the clinical review and no treatment. If a decision model is produced, expected value of perfect information (EVPI) analyses can be conducted and will be used to determine the need for further research, identify the research questions critical to decision-making, and help inform the design of future studies.
Chapter 3  Assessment of the clinical effectiveness and cost-effectiveness evidence

Methods for reviewing the evidence

The review was conducted systematically following the general principles recommended in the Centre for Reviews and Dissemination’s guidance for undertaking reviews in health care\(^62\) and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.\(^63\)–\(^65\)

Identification of studies

The screening of titles and abstracts was conducted by two independent reviewers. All potentially relevant studies were retrieved where available, and two independent reviewers applied the inclusion criteria to the full papers. Disagreements were resolved by team discussion. Where consensus could not be reached at the title and abstract stage, the full paper was ordered. Inclusion was not restricted by language or date of publication. Abstracts were included if no additional information was available and there were sufficient outcome data to extract.

Search strategy

The aim of the literature searches was to systematically identify studies on the effectiveness, test accuracy, test reliability, test reproducibility and cost-effectiveness of bone turnover markers in people being treated for osteoporosis. Search terms were identified by scanning key papers identified at the beginning of the project, through discussion with the review team and through the use of database thesauri. The creation of the search strategy was an iterative process originally using the MEDLINE database and then adapted as appropriate to the other sources searched.

The base search strategy included the following components:

1. bone turnover marker terms
   AND
2. osteoporosis terms
   AND
3. intervention terms.

Sources of information were identified by an information specialist with input from the project team. The following databases were searched without language or date restrictions to identify primary studies, relevant reviews and economic studies:

- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (via EBSCO\textit{host} 1982 to 4 March 2012)
- The Cochrane Library (Issue 2 of 12 February 2012), which includes: Cochrane Database of Systematic Reviews Database of Abstracts of Reviews of Effects (DARE) Cochrane Central Register of Controlled Trials NHS Economic Evaluation Database (NHS EED) Health Technology Assessment Database
- Conference Proceedings Citation Index – Science (via Web of Knowledge 1990 to March 2012)
- EconLit (via Ovid\textit{SP} 1961 to February 2012)
Ongoing research

Ongoing studies were identified from the following databases:

- ClinicalTrials.gov (via website at www.clinicaltrials.gov to March 2012)
- ControlledTrials.com (via website at http://controlled-trials.com to March 2012)
- Paid Clinical Trials (via website at www.paidclinicaltrials.org to May 2012).

Other sources

The reference lists of included papers and relevant reviews were assessed for additional relevant studies. Where necessary, authors of eligible studies were contacted for further information and experts in the field were contacted to see if they had access to further material. We also hand-searched (in May 2012) the contents pages for the previous 12 months of three relevant journals that were the source of a large proportion of identified studies (Osteoporosis International, Journal of Bone and Mineral Research and Bone) as these recent issues may be poorly indexed in the electronic databases.

The websites of the following organisations were also searched in May 2012 for information on relevant trials and other research:

- Eli Lilly and Company: www.lilly.com/
- GlaxoSmithKline: www.gsk.com/
- Novartis: www.novartis.com/
- Nycomed: www.nycomed.com/
- Procter & Gamble: www.pg.com/
- US Food and Drug Administration (FDA): www.fda.gov/
  - FDA website search included specific searches of:
    - Office for Women’s Health (OWH): Research Science Program Awards: osteoporosis section
    - Publications based on OWH projects: osteoporosis section
    - Medical devices section.

The total number of records found after deduplication was 4002. Records were managed within an EndNote library (EndNote version X3, Thomson Reuters, CA, USA). The full search strategies for each database searched are provided in Appendix 1.

Inclusion and exclusion criteria

Index tests being evaluated

The review evaluated four bone turnover marker tests, two serum bone formation markers (sP1NP and sBALP), and two bone resorption markers that can be measured in either the serum or urine (s/uCTX and s/uNTX).
Population
Studies eligible for inclusion were those in adults (> 18 years of age) either:

- receiving any of bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for the secondary prevention of osteoporotic fractures, regardless of the baseline pathology; or
- in any high-risk group being treated with any of bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for the primary prevention of osteoporotic fractures.

Study designs

Effectiveness
Randomised controlled trials (RCTs) of any size, where patients are randomised to a standard monitoring regimen (with or without DXA) or to a standard monitoring regimen with additional bone turnover marker monitoring. Studies reporting the impact of bone turnover marker test results on the decision-making process for management of osteoporosis, that also reported the subsequent rate of fracture in the population being assessed, were also sought (‘Decision studies’). Studies assessing the effectiveness of treatments for osteoporosis using changes in bone turnover markers solely as an outcome were excluded.

Test accuracy
Studies comparing the results of bone turnover marker tests with the results of bone biopsy or a composite reference standard of DXA and subsequent fracture outcome were included. Given the nature of the review question, we believe it unlikely that such studies would be available, so in addition we included prospective studies that measured the association between bone turnover and bone density, biopsy results and/or fracture rates, and that reported a correlation coefficient for this association. Prospective studies that evaluated changes in bone turnover markers in patients receiving one of the specified osteoporosis treatments, that provided sufficient data to produce a measure of the risk of fracture, or that reported the results of multivariate regression analyses in which a bone turnover marker of interest is an independent variable, were also eligible for inclusion. Prognostic studies using a bone turnover marker to identify patients at risk of osteoporosis and fracture at baseline, prior to commencing treatment, were excluded, as were studies that included fewer than 20 patients, meeting the population inclusion criteria, in analyses of outcomes applicable to this review.

Reliability and reproducibility
Prospective controlled studies of serial bone turnover marker measurements that reported a measure of within- and/or between-patient variability in patients receiving a treatment being evaluated in this review were included. Inclusion was restricted to studies that included at least 20 patients in at least one analysis of interest.

Economic evaluation
Full economic evaluations meeting the population and intervention inclusion criteria. A full economic evaluation was defined as any study in which a comparison of two or more relevant alternatives was undertaken with costs and outcomes examined separately for each alternative.

Outcomes

Effectiveness
Randomised controlled trials and decision studies reporting either change in patient management strategies, the incidence of fracture and/or treatment adherence rates were included.

Test accuracy
Studies had to report either:

- estimates of diagnostic accuracy, or sufficient data for these to be calculated
- a correlation coefficient, or sufficient data for this to be calculated, for the association between a bone turnover marker and bone density and/or the incidence of fracture
the risk/incidence of fracture associated with the bone turnover marker test results

at least a p-value for a bone turnover marker of interest that is used as an independent variable in a multivariate regression.

**Reliability and reproducibility**

Studies reporting a measure for intra- and/or interpatient variability in bone turnover marker test results were included.

**Economic evaluation**

Study inclusion was not restricted by outcome.

**Data extraction strategy**

Clinical data extraction was conducted by one reviewer using a standardised data extraction form and checked by a second reviewer. Discrepancies were resolved by discussion, with involvement of a third reviewer when necessary. Attempts were made to contact authors for missing data. Data from multiple publications of the same study were extracted and reported as a single study. Where applicable and available, extraction included data on study details (e.g. study/EndNote identifier, author, year, country, setting, number of participants and duration of follow-up), patient characteristics (e.g. age, gender, duration of osteoporosis, risk group, concomitant renal/liver disease; baseline bone turnover marker levels and BMD), details of intervention (serum or urine; sample collection details; pre-sampling preparations/restrictions; sample storage details; assay used; adjustments for creatinine excretion for urinary markers; delay between sample collection and assay; single/serial measures; intra- and interassay coefficients of variation; value for least significant change), study quality, and reported outcomes as specified above.

Economic data extraction was planned on the comparators, study population, main analytic approaches (e.g. patient-level analysis/decision-analytic modelling), primary outcome specified for the economic analysis, details of adjustment for QoL, direct costs (medical and non-medical) and productivity costs, estimates of incremental cost-effectiveness and approaches to quantifying decision uncertainty (e.g. deterministic/probabilistic sensitivity analysis).

**Critical appraisal strategy**

The quality of the individual studies was assessed at the study level by one reviewer and independently checked by a second; disagreements were resolved by consensus. The quality of included studies was assessed using relevant criteria suitable for the study design selected from standard checklists for RCTs, observational studies and economic evaluations; topic-specific quality issues were incorporated where necessary (see Appendix 2 for details and guidance for completion).

**Methods of data synthesis**

Key study characteristics, patient outcomes and study quality were summarised in a narrative and tables. Meta-analyses suitable to the clinical data extracted were planned to estimate a summary measure of effect when sufficient numbers of comparable studies were available for an outcome. Given the substantial heterogeneity across the studies, this was not possible. It was also not possible to investigate the potential sources of heterogeneity that were specified in the protocol, as insufficient numbers of studies similar in other population, intervention and methodological characteristics were identified. The analyses planned were:

- investigation of potential subgroups of interest where sufficient data are available; for example, post-menopausal women (overall and for specific age ranges if data are available), elderly, skeletal site (hip, spine or wrist), and glucocorticoid-induced osteoporosis
- sensitivity analyses conducted, where appropriate, to investigate potential sources of heterogeneity such as study quality, and differences in sample acquisition, storage and assay methods.

A narrative synthesis of the methods and results of the cost-effectiveness studies was planned.
Results of the evaluation of clinical effectiveness

Quantity and overall quality of research available

As a result of the electronic and hand searching, 4002 papers were identified for initial screening. Of these, 437 were retrieved as full papers; 35 were reviews that underwent screening of their bibliographies. Sixteen of the full papers were published in languages other than English. Nineteen studies had duplicate publications; where this was the case one paper was allocated as the primary publication (full paper if the duplicate was an abstract or letter; the published paper if the duplicate comprised data from the manufacturer’s online trials database, the most recent publication, or the paper published in English) and used as the citation for the study throughout the review. Relevant data were extracted from all publications where applicable. After full-paper screening, 42 studies (across 70 publications) met the inclusion criteria for the review of clinical effectiveness; the flow of studies through the review is given in Figure 1 (some of the studies were excluded for more than one reason). A list of excluded studies with the reasons for exclusion is given in Appendix 3.

The majority of the studies were conducted primarily, or entirely, in post-menopausal women; only four studies reported including men. Where reported, the mean age ranged from 56.1 to 73.9 years.

---

**FIGURE 1** Flow of studies through the review.
The definition used for the diagnosis of osteoporosis differed across studies; the definition was not provided in nine studies.106,133,135,143,145,147–150 Most of the included studies were small, with the number of participants ranging from 22 to 3105 where reported (two studies did not report the number recruited40,151); 21 studies had fewer than 100 participants.41,43,44,58,99,106,135,136,145,147,150,152–160 A summary of study characteristics is given in Table 1; full data extraction tables are provided in Appendix 4.

### Table 1 Summary of study characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatments</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong (2007),145 UK</td>
<td>Definition OP: NR</td>
<td>sCTX; no details</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate or risedronate; no details</td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td>n = 46; n with OP = 46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 6; n PMW = NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td>Bauer (2004),139 USA/Canada</td>
<td>Definition OP: T-score ≤ −2.5; vertebral fracture</td>
<td></td>
</tr>
<tr>
<td>Started 1992</td>
<td>Alendronate 5 mg/day increased to 10 mg/day at second annual visit for 2 years</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 3105; n with OP = 3105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 3105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td>Bjarnason (2001),151 multinational</td>
<td>Definition OP: T-score LS/FN ≤ −2.5</td>
<td></td>
</tr>
<tr>
<td>Dates NR</td>
<td>Raloxifene 60 or 120 mg/day</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = NR; n male = 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td>Blumsohn (2011),42 Western Europe</td>
<td>Definition OP: T-score LS/hip/FN ≤ −2.5 + ≥ 1 OP fracture past 3 years</td>
<td></td>
</tr>
<tr>
<td>Dates NR</td>
<td>Teriparatide 20 µg/day for 1 or 2 years</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 758; n with OP = 758</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 758</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 69.8 years</td>
<td></td>
</tr>
<tr>
<td>Bruyere (2010),161 Western Europe (RCTs multinational)</td>
<td>Definition OP: T-score ≤ −2.5 + ≥ 1 risk factor</td>
<td></td>
</tr>
<tr>
<td>Dates NR</td>
<td>Strontium ranelate 2 g/day for NR</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 2373; n with OP = 2373</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 2373</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 73.9 years</td>
<td></td>
</tr>
<tr>
<td>Burshell (2010),14 USA/Canada</td>
<td>Definition OP: T-score LS/hip ≤ −1.0 + ≥ 1 OP fracture; T-score at LS/hip ≤ −2.0</td>
<td></td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 10 mg/day for at least 18 months</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 77; n with OP = 77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 17; n PMW = 50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 60.6 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teriparatide 20 µg/day for at least 18 months</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 80; n with OP = 80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 13; n PMW = 41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 56.1 years</td>
<td></td>
</tr>
<tr>
<td>Chen (2005),140 multinational</td>
<td>Definition OP: T-score LS/hip ≤ −1.0 + ≥ 1 OP fracture; one moderate or two mild vertebral fractures</td>
<td></td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CLINICAL EFFECTIVENESS AND COST-EFFECTIVENESS EVIDENCE
### TABLE 1 Summary of study characteristics (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatments</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clowes (2003), UK</td>
<td>Teriparatide 20 µg/day for median 19 months n = 541; n with OP = 541 n male = 0; n PMW = 541 Mean age: NR</td>
<td>sBALP; IRMA uNTX; ELISA Baseline; 1, 3, 6, 12 months; study end</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Teriparatide 40 µg/day for median 19 months n = 552; n with OP = 552 n male = 0; n PMW = 552 Mean age: NR</td>
<td>sP1NP, assay method NR sCTX; ECL Baseline; 12, 18 months</td>
</tr>
<tr>
<td>Delmas (2007), multinational</td>
<td>Definition OP: NR</td>
<td>DXA; FN; LS</td>
</tr>
<tr>
<td>Dates 1999 to 2002</td>
<td>Raloxifene 60 mg/day for NR n = 22; n with OP = 22 n male = NR; n PMW = NR Mean age: NR</td>
<td>BM feedback; 13, 25 weeks n = 1189 Mean age: 71.1</td>
</tr>
<tr>
<td>Full published paper</td>
<td>Definition OP: T-score LS/hip ≤ −2.5</td>
<td>uNTX; ELISA</td>
</tr>
<tr>
<td>Delmas (2007), multinational</td>
<td>Risedronate 5 mg/day for 1 year n = 2382; n with OP = 2382 n male = 0; n PMW = 2382 Mean age: NR</td>
<td>No BM feedback n = 1113 Mean age: 71.5 years</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Zoledronate 5 mg/year for 3 years n = NR; n male = 0 Mean age: NR</td>
<td>sCTX; ECL sP1NP, ECL sBALP; ELISA</td>
</tr>
<tr>
<td>Full published paper</td>
<td>Definition OP: T-score ≤ −1.5 + one moderate or two mild vertebral fractures; T-score LS/hip ≤ −2.5</td>
<td>BM feedback; 13, 12, 18 months; 1, 3, 6, 12 months after third infusion</td>
</tr>
<tr>
<td>Delmas (2009), multinational</td>
<td>Teriparatide 20 or 40 µg/day for 17 to 22 months n = 36; n with OP = 36 n male = 0; n PMW = 36 Mean age: 67.9 years</td>
<td>uNTX; ELISA Baseline; 6, 12, 24, 36 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Definition OP: one moderate or two mild vertebral fractures; T-score LS/hip ≤ −1.0 + ≥ 1 OP fracture</td>
<td>sCTX; ECL sP1NP, ECL sBALP; ELISA</td>
</tr>
<tr>
<td>Full published paper</td>
<td>Amlodipine 10 mg/day or risedronate 5 mg/day n = 37; n with OP = 37 n male = 0; n PMW = 37 Mean age: 69 years</td>
<td>DXA; FN Baseline; 2, 6, 12 months</td>
</tr>
<tr>
<td>Dobnig (2006), western Europe</td>
<td>Definition OP: T-score LS/hip ≤ −2.5</td>
<td>DXA; FN; LS (L1–L4)</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Teriparatide 20 or 40 µg/day for 17 to 22 months n = 36; n with OP = 36 n male = 0; n PMW = 36 Mean age: 67.9 years</td>
<td>DXA; FN Baseline; 2, 6, 12 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td>Zoledronate 5 mg/year for 3 years n = NR; n male = 0 Mean age: NR</td>
<td>sCTX; ECL sP1NP, ECL sBALP; ELISA</td>
</tr>
<tr>
<td>Eastell (2003), multinational</td>
<td>Definition OP: two vertebral fractures; one vertebral fracture and T-score &lt; −2</td>
<td>uNTX; chemiluminescence uCTX; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Risedronate 5 mg/day for 3 years n = 358; n with OP = 358 n male = 0; n PMW = 358 Mean age: 70 years</td>
<td>DXA; FN; LS (L1–L4) Baseline; 12, 36 months</td>
</tr>
</tbody>
</table>

**continued**
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatments</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastell (2011), 43 multinational</td>
<td>Definition OP: T-score LS/hip ≤−2.5</td>
<td>sCTX; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Denosumab 60 mg every 6 months for 3 years</td>
<td>sP1NP; RIA</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 96; n with OP = 96</td>
<td>sP1NP (intact); RIA</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 96</td>
<td>sP1NP (total); ECL</td>
</tr>
<tr>
<td></td>
<td>Mean age: 72.3 years</td>
<td>sCTX; ECL</td>
</tr>
<tr>
<td>Garnero (2008), 41 USA/Canada</td>
<td>Definition OP: T-score LS/hip ≤−2.0 + 1 risk factor; T-score LS/hip ≤−2.5</td>
<td>Baseline; 1, 6, 12, 24, 36 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 10 mg/day for 3 months</td>
<td>DXA; hip; LS</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 60; n with OP = 60</td>
<td>Baseline; 12, 24, 36 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 70.7 years</td>
<td></td>
</tr>
<tr>
<td>Heaney (2011), 162 USA/Canada</td>
<td>Definition OP: T-score LS/hip ≤−1.0 + ≥1 OP fracture</td>
<td>Baseline; 3, 6, 12 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Teriparatide 20 µg/day</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 203; n with OP = 203</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 203</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 70 years</td>
<td></td>
</tr>
<tr>
<td>Hochberg (2010), 163 USA/Canada</td>
<td>Definition OP: T-score LS/hip ≤−2.5</td>
<td>Baseline; 3, 6, 12 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Ibandronate 150 mg monthly for 1 year</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 323; n with OP = 323</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 323</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 65.8 years</td>
<td></td>
</tr>
<tr>
<td>Imai (2009), 136 Asia</td>
<td>Definition OP: LS BMD ≤70% YAM; vertebral fracture</td>
<td>Baseline; 3 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 5 mg/day for 1 year</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 37; n with OP = 37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 76.5 years</td>
<td></td>
</tr>
<tr>
<td>Ishijima (2009), 154 Asia</td>
<td>Definition OP: LS BMD ≤70% YAM; LS BMD ≤80% YAM + ≥1 fracture</td>
<td>Baseline; 6 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 5 mg/day for 6 months</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 45; n with OP = 45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 70.2 years</td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2004), 155 Asia</td>
<td>Definition OP: LS BMD ≤70% YAM; LS BMD ≤80% YAM + ≥1 fracture</td>
<td>Baseline; 6 months</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 5 mg/day for 12 months</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 85; n with OP = 85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 72.2 years</td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2005), 131 Asia</td>
<td>Definition OP: LS BMD ≤70% YAM; LS BMD ≤80% YAM + ≥1 fracture</td>
<td>Baseline; 6, 12 months</td>
</tr>
<tr>
<td>2002 to 2004</td>
<td>Alendronate 5 mg/day for 1 year</td>
<td></td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 132; n with OP = 132</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 132</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 71.9 years</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Population and treatments</td>
<td>Interventions</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
</tr>
<tr>
<td>Kim (2005),[^44] Asia</td>
<td><strong>Definition of OP:</strong> T-score ≤ 2.5 SD below normal mean for Korean PMW at LS</td>
<td>uNTX; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>Baseline; 3, 6 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Alendronate 10 mg/day for 1 year</strong></td>
<td>DXA; FN; LS (L1–L4)</td>
</tr>
<tr>
<td></td>
<td>n = 50; n with OP = 50</td>
<td>Baseline; 12 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 60.3 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline; 3, 6 months</strong></td>
<td></td>
</tr>
<tr>
<td>Kitatani (2003),[^156] Asia</td>
<td><strong>Definition of OP:</strong> LS BMD ≤ 70% YAM</td>
<td>sBALP; EIA</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>Baseline; 3, 6, 12 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Etidronate; 200 mg/day for 98 weeks; 2 weeks with drug followed by 10 weeks without</strong></td>
<td>DXA; LS (L2–L4)</td>
</tr>
<tr>
<td></td>
<td>n = 32; n with OP = 32</td>
<td>Baseline; 6, 12, 18, 24 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 63.3 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Etidronate; 400 mg/day for 98 weeks; 2 weeks with drug followed by 10 weeks without</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 31; n with OP = 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 64.8 years</td>
<td></td>
</tr>
<tr>
<td>Kung (2009),[^133] Asia</td>
<td><strong>Definition OP:</strong> NR</td>
<td>sCTX; assay method NR</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>Baseline; 3, 6 months</td>
</tr>
<tr>
<td>Manufacturer’s trial</td>
<td><strong>Ibandronate 150 mg monthly for 12 months</strong></td>
<td>BM feedback; 3 months</td>
</tr>
<tr>
<td>database/full paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 596; n with OP = 596</td>
<td>n = 300</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 596</td>
<td>Mean age: 66.3 years</td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyd (1998),[^157] UK</td>
<td><strong>Definition OP:</strong> T-score LS/FN ≤ –2.5</td>
<td>sBALP-i; IRMA</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>sBALP-E; ICEA</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Alendronate 10 mg/day for 1 year</strong></td>
<td>Baseline; 3 months</td>
</tr>
<tr>
<td></td>
<td>n = 35; n with OP = 35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median age: 67 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyd (1999),[^158] UK</td>
<td><strong>Definition OP:</strong> T-score LS/FN ≤ –2.5</td>
<td>sCTX; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>Baseline; 3, 6 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Alendronate 10 mg/day for 1 year</strong></td>
<td>DMX; FN; spine</td>
</tr>
<tr>
<td></td>
<td>n = 30; n with OP = 30</td>
<td>Baseline; 12 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane (2000),[^159] USA/Canada</td>
<td><strong>Definition OP:</strong> T-score FN ≤ –2.5, T-score LS/hip ≤ –2.5</td>
<td>sBALP; EIA</td>
</tr>
<tr>
<td>Dates NR</td>
<td></td>
<td>Baseline; 1, 3, 6, 9, 18, 24 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Teriparatide 40 µg/day</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 28; n with OP = 28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majima (2008),[^160] Asia</td>
<td><strong>Definition OP:</strong> T-score LS/hip ≤ –2.5</td>
<td>sNTX; ELISA</td>
</tr>
<tr>
<td>2004 to 2007</td>
<td></td>
<td>Baseline; 3, 6, 12 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td><strong>Raloxifene 60 mg/day for 12 months</strong></td>
<td>DXA; FN; LS (trochanter, radius; Ward’s triangle)</td>
</tr>
<tr>
<td></td>
<td>n = 63; n with OP = 63</td>
<td>Baseline; 6, 12, 12 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 70.5 years</td>
<td></td>
</tr>
</tbody>
</table>

**continued**
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatments</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masaryk (2002),99 eastern Europe</td>
<td>Definition OP: T-score LS/hip ≤ –2.5</td>
<td>uNTX; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Alendronate 10 mg/day for 12 months</td>
<td>Baseline; 3 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 50; n with OP = 50</td>
<td>DXA; FN; LS (L2-L4); TB; trochanter</td>
</tr>
<tr>
<td></td>
<td>n male = 50; n PMW = 50</td>
<td>Baseline; 12 months</td>
</tr>
<tr>
<td></td>
<td>Mean age: 64.2 years</td>
<td></td>
</tr>
<tr>
<td>Miller (2008),38 multinational</td>
<td>Definition OP: T-score LS/hip ≤ –2.5 + ≥ 1 OP fracture</td>
<td>sP1NP; ECL</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Teriparatide 20 µg/day for 1 year</td>
<td>Baseline; 0.5, 1, 2, 3, 4, 5, 6, 9, 12 months</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 317; n with OP = 317</td>
<td>DXA; hip; LS</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 317</td>
<td>Baseline; 6, 12 months</td>
</tr>
<tr>
<td>Moro-Alvarez (2010),135 western Europe</td>
<td>Definition OP: NR</td>
<td>sCTX; ECL</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Strontium ranelate 2 g/day for 12 to 24 months</td>
<td>sP1NP; RIA</td>
</tr>
<tr>
<td>Abstract</td>
<td>n = 66; n with OP = 66</td>
<td>Baseline; 12, 24 months</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 66</td>
<td>DXA; FN; LS (L2-L4); total hip</td>
</tr>
<tr>
<td></td>
<td>Mean age: 68 years</td>
<td>Baseline; 24 months</td>
</tr>
<tr>
<td>Reginster (2004),132 multinational</td>
<td>Definition OP: T-score LS/hip ≤ –2.5; vertebral fracture</td>
<td>sBALP; IRMA</td>
</tr>
<tr>
<td>Dates NR</td>
<td>Raloxifene 60 mg/day for up to 3 years</td>
<td>uCTX; ELISA</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n = 347; n with OP = 347</td>
<td>sP1NP; RIA</td>
</tr>
<tr>
<td></td>
<td>n male = 0; n PMW = 347</td>
<td>Baseline; 6, 12, 24, 36 months</td>
</tr>
<tr>
<td></td>
<td>Mean age: 68.2 years</td>
<td></td>
</tr>
<tr>
<td>Reyes-Garcia (2010),58 western Europe</td>
<td>Raloxifene 120 mg/day for up to 3 years</td>
<td>sBALP; ELISA</td>
</tr>
<tr>
<td>Dates NR</td>
<td>n = 254; n with OP = 254</td>
<td>sCTX; ECL</td>
</tr>
<tr>
<td>Full published paper</td>
<td>n male = 0; n PMW = 254</td>
<td>Baseline; 3, 6, 12 months</td>
</tr>
<tr>
<td>Roche (2007),143 South America</td>
<td>Definition OP: NR</td>
<td>sCTX; assay method NR</td>
</tr>
<tr>
<td>2006 to 2007</td>
<td>Alendronate 70 mg/week for 1 year</td>
<td>Baseline; 3 months (feedback arm); 6 months</td>
</tr>
<tr>
<td>Manufacturer’s trial</td>
<td>n = 46; n with OP = 46</td>
<td>BM feedback; 3 months</td>
</tr>
<tr>
<td>database</td>
<td>n male = 0; n PMW = 46</td>
<td>n = NR</td>
</tr>
<tr>
<td></td>
<td>Mean age: 64.7 years</td>
<td>Mean age: NR</td>
</tr>
<tr>
<td></td>
<td>No BM feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = NR</td>
<td></td>
</tr>
<tr>
<td>Roche (2009),148 multinational</td>
<td>Definition OP: NR</td>
<td>sCTX; assay method NR</td>
</tr>
<tr>
<td>2007 to 2008</td>
<td>Ibandronate 150 mg monthly for 6 months</td>
<td>Baseline; 1.5 months</td>
</tr>
<tr>
<td>Manufacturer’s trial</td>
<td>n = 585; n with OP = 585</td>
<td>BM feedback; approx. 2 months</td>
</tr>
<tr>
<td>database</td>
<td>n male = 0; n PMW = 585</td>
<td>No baseline details</td>
</tr>
<tr>
<td></td>
<td>Mean age: NR</td>
<td>No BM feedback</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No baseline details</td>
</tr>
</tbody>
</table>
### TABLE 1 Summary of study characteristics (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatments</th>
<th>Interventions</th>
</tr>
</thead>
</table>
| Roche (2009),149 western Europe | Ibandronate 150 mg monthly for 12 months  
  *n* = 596; *n* with OP = 596  
  *n* male = 0; *n* PMW = 596  
  Mean age: NR  
  Dates NR | BM feedback; after 5-week test  
  *n* = 250  
  Mean age: NR  
  Manufacturer’s trial database | sCTX; Assay method NR  
  Baseline: 5 weeks; 3, 6, 12 months |
| Sarkar (2004)164 Multinational Dates NR | Raloxifene 60 or 120 mg/day  
  *n* = 1650; *n* with OP = 1650  
  *n* male = 0; *n* PMW = 1650  
  Mean age: 67.3 years  
  Dates NR | DXA; FN; LS (L2–L4)  
  Annually  
  Full published paper | sBALP; IRMA  
  Baseline: 6, 12, 24, 36 months |
| Shiraki (2011),142 Asia 2000 to 2009 | Alendronate 5 mg/day or 35 mg/week and risedronate 2.5 mg/day or 17.5 mg/week for mean 3.2 years  
  *n* = 251; *n* with OP = 251  
  *n* male = 0; *n* PMW = 251  
  Mean age: 70.5 years  
  Dates NR | DXA; LS  
  Baseline: every 6 months  
  Full published paper | uNTX; ELISA  
  sBAP; EIA  
  Baseline: 6 months intervals; study end |
| Siddiqi (2010),106 UK Dates NR Abstract | Teriparatide for 18 months; dose NR  
  *n* = 28; *n* with OP = 28  
  *n* male = 0; *n* PMW = NR  
  Mean age: 74 years  
  Dates NR | sP1NP; assay method NR  
  Baseline: 3 months  
  Abstract | DXA; spine  
  Baseline: 18 months |
| Stepan (2008),150 multinational Dates NR Abstract | Teriparatide 20 µg/day for 24 months  
  *n* = 66; *n* with OP = 66  
  *n* male = 0; *n* PMW = 66  
  Mean age: 68 years  
  Dates NR | sCTX; assay method NR  
  sP1NP; assay method NR  
  Baseline: 1, 3, 6, 12, 24 months  
  Abstract | Biopsy; iliac crest  
  24 months |
| Tsujimoto (2011),146 Asia Dates NR | Teriparatide 20 µg/day for 12 months  
  *n* = 136; *n* with OP = 136  
  *n* male = 9; *n* PMW = 127  
  Mean age: 69.2 years  
  Dates NR | sP1NP; RIA  
  sBAP; ostease (variant NR)  
  sCTX; ELISA  
  Baseline: 1 month; 3, 6, 12 months  
  Full published paper | DXA; FN; LS (L2–L4)  
  Baseline: 3, 6, 12 months |
| Watts (2001),165 multinational Dates NR | Alendronate 10 mg/day for at least 1 year  
  *n* = 180; *n* with OP = 180  
  *n* male = 0; *n* PMW = 180  
  Mean age: NR  
  Dates NR | DXA; FN; LS (L1–L4); TB  
  Baseline: 3, 6, 12, 18, 24, 36 months  
  Full published paper | sBAP; EIA  
  Baseline: 3, 6, 12 months |

ECL, electrochemiluminescence; EIA, enzyme immunoassay; FN, femoral neck; ICEA, immunocapture enzymatic assay; intact P1NP, measurement of the trimetric forms only; L1, L2, L4, lumbar vertebrae 1, 2, 4; LS, Lumbar spine; *n*, number of patients; NR, not reported; OP, osteoporosis; PMW, post-menopausal women; RIA, radioimmunoassay; total P1NP, measurement of the mono- and trimetric forms; TB, total body; YAM, young adult mean.
Quantity and quality of the included randomised controlled trials

Of the 42 included studies, five were RCTs. Of these, one reported using an adequate method for randomisation and allocation concealment; the methods were not reported for the other four RCTs. As a result, the appropriateness of the control group in the four RCTs was also deemed to be unclear; however, given the limited information available, it is unlikely that the control group was not selected from the same population as the intervention group. Only one RCT recruited a representative osteoporotic patient population; the others were conducted in a selected subgroup of post-menopausal women, restricted either by ethnicity or by age. Baseline comparability could be assessed in only two trials; groups were comparable. Although blinding of patients and care givers is not feasible for these types of interventions, blinding of outcome assessors is; none of the RCTs reported blinding outcome assessors. Descriptions of the intervention details were considered adequate to allow replication in two RCTs. None of the five RCTs reported the characteristics of the patients lost to follow-up, and only three reported reasons for the losses. Three of the five reported using an intention-to-treat (ITT) analysis, and one used ITT for some analyses, and one reported excluding patients who were randomised but did not return electronic monitors from the ITT population. The imputation methods used for missing data were not reported for any of the RCTs. Four of the RCTs had a period of follow-up less than the 1 year considered by the authors of this review to be the minimum duration required to identify changes in treatment strategies and subsequent fracture risk; however, none of the RCTs assessed either of these outcomes. Given the limitations of the included RCTs, all were considered to be at a high or uncertain risk of bias, and therefore of low quality. The full results of the quality assessment and the guidance used for its completion are given in Appendix 2.

Quantity and quality of the included non-randomised studies

Of the 37 non-randomised studies, 21 were cohorts derived from the treatment arms of RCTs that compared a treatment regimen with either placebo or an alternative treatment. The cohorts were derived either directly from reports of the RCT or from a paper reporting a post hoc analysis of the RCT or from a post hoc analysis of a subgroup of patients from the RCT. These derived cohorts were assessed for quality as cohort studies rather than RCTs, as this is the manner in which they were used in the review.

The remaining 16 studies were 14 uncontrolled cohort studies and two controlled cohort studies; one compared two groups with different prior treatment regimens, and the other treated one group of women with HRT and the other with alendronate, although the groups seem to have been established after recruitment. Single cohorts were derived from both controlled cohort studies. In the only truly controlled cohort study, the recruitment of the control group was considered appropriate and the groups comparable at baseline. Of the 16 studies that were designed as cohort studies, only four appeared to use consecutive recruitment (though this was usually inferred, not explicitly stated), and one was not consecutive, and recruitment was unclear in 11. Across the 37 non-randomised studies, 16 recruited a representative population, and in the remaining four, it was unclear whether or not the population was representative. Ten studies had no loss to follow-up for the analyses they conducted (some were post hoc analyses of the specific group of patients with the required outcome measures). The controlled cohort study used a ‘modified’ ITT analysis; seven patients were excluded from the ITT population, and last observation carried forward was used for the imputation of missing data for those in the modified ITT population. The reasons were given for losses to follow-up in a further 11 studies.
Confounders were clearly identified and described in five studies. Although six studies adjusted for confounding factors in multiple regression analyses, only two adjusted for all the confounders considered important by the review authors (age, gender, prior fracture, baseline BMD, and BMI). Twenty-seven of the studies had a minimum of 1-year follow-up in all patients. The assessment of the reporting of adverse events of bone turnover markers was not assessed for these studies as their focus was not the assessment of the effectiveness or safety of bone turnover markers. A summary of the results of the quality assessment is given in Figure 2.

Given the limitations of the non-randomised studies included, all were considered to be at a high or unclear risk of bias, and therefore the overall quality of each of the studies was considered to be low. The low quality assigned to the cohorts derived from RCTs is not necessarily a reflection of the quality of the original RCT; this classification was primarily driven by the post hoc nature of the selection of patients in the paper reporting the outcomes of interest for this review, which can introduce bias. The full results of the quality assessment and the guidance used for its completion, including the criteria on which the overall quality was primarily based upon, are given in Appendix 2.

FIGURE 2 Summary of the quality of the non-randomised studies.
**Assessment of clinical effectiveness**

Five RCTs56,133,143,148,149 and one post hoc analysis of a subgroup of patients41 from a previous RCT166 evaluated the effectiveness of feedback of bone turnover marker results on QoL and/or adherence, compliance, or persistence in post-menopausal women. These terms were defined variably across the studies, particularly the proportion of medication taken which was the definition used across the studies for adherence, compliance and persistence. We decided to standardise the definitions by adopting the system used by Delmas et al.56: persistence is the time until discontinuation of medication, compliance is the proportion of medication taken, and adherence is a combination of persistence and compliance. None of the RCTs reported on the impact of bone turnover marker monitoring on treatment management.

**Adherence**

One study reported on adherence and found a significantly greater average daily proportion of patients who were persistent and compliant to 5 mg/day risedronate in those receiving feedback after the first reinforcement visit at 13 weeks ($p = 0.01$).56 The decrease in adherence observed over time was attributed to the increasing number of patients who did not persist with treatment.

**Compliance**

Five studies reported on compliance (Table 2).41,133,143,148,149 The rate of compliance with bisphosphonate therapy, even in the no-feedback arms of the trials, was very high and unlikely to be representative of clinical practice. Therefore, with such a high rate of baseline compliance, the capacity for feedback of bone turnover marker results to impact on the compliance is limited and there seems to be little difference between the feedback and no-feedback arms of the trials. Where odds ratios (ORs) could be calculated, there was no significant difference between the feedback and no-feedback arms; this could not be calculated for most of the studies owing to insufficient data (see Table 2).

**Persistence**

One study reported 77% of patients persisting with 5 mg/day risedronate in those receiving feedback of uNTX results, and 80% in those that were not; this was a high baseline rate that may not be representative of clinical practice.56 When adjusted for compliance, there was no significant impact of uNTX on persistence in the no-feedback arm ($p = 0.71$); feedback of uNTX results significantly affected persistence in the feedback group ($p = 0.0029$). Overall, there was a significant impact of feedback of uNTX results on discontinuation ($p = 0.017$); where the message given to the patient was a good uNTX response (> 30% decrease), the hazard ratio (HR) for discontinuation was 0.71 [95% confidence interval (CI) 0.53 to 0.95]. Where the message given was a poor uNTX response, the HR for discontinuation was 2.22 (95% CI 1.27 to 3.89). Where the message given was that uNTX was stable, there was no significant difference in discontinuation between those receiving feedback and those who were not.56

**Quality of life**

Two studies reported the results of the Osteoporosis Patient Perception Survey (OPPS) QoL questionnaire.133,148 One reported statistically significant differences of at least 3.8% favouring feedback for all domains and the composite score ($p \leq 0.021$) in women aged between 55 and 85 years, with the exception of the motivation domain ($p > 0.05$).148 The mean scores reported in the second study are given in Table 3; there were significant increases in feeling informed, satisfaction and the overall composite score.133 Although statistically significant, it is unclear whether or not the small absolute changes are clinically significant.
<table>
<thead>
<tr>
<th>Trial</th>
<th>Population; treatment</th>
<th>BM timepoint</th>
<th>Definition; treatment period</th>
<th>BM feedback</th>
<th>Compliant</th>
<th>Non-compliant</th>
<th>OR: 95% CI</th>
<th>BM feedback</th>
<th>Compliant</th>
<th>Non-compliant</th>
<th>OR: 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kung (2009)</td>
<td>PMW ≤ 85 years; 150 mg/monthlyibandronate</td>
<td>≤ 3 months</td>
<td>At least five of six monthly doses within 1- to 21-day windows; 6 months</td>
<td>No feedback</td>
<td>n = 274</td>
<td>n = 12</td>
<td>96% (4%)</td>
<td>22 (7%)</td>
<td>1.93 (0.94 to 3.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roche (2009)</td>
<td>PMW ≤ 55 years; 150 mg/monthlyibandronate</td>
<td>1.5 months</td>
<td>At least 10 of 12 monthly doses within 1- to 21-day windows; 12 months</td>
<td>Feedback</td>
<td>65.7%</td>
<td>11%</td>
<td>65.7%</td>
<td>29.8%</td>
<td>No significant difference between feedback and no feedback arms (p = 0.132)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-feedback</td>
<td>34.3%</td>
<td>6%</td>
<td>34.3%</td>
<td>29.8%</td>
<td>OR: 0.98 (0.67 to 1.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garnero (2008)</td>
<td>PMW 55–85 years; 10 mg/day alendronate</td>
<td>3 months</td>
<td>80% of pills taken; 3 months</td>
<td>No feedback</td>
<td>n = 49</td>
<td>82%</td>
<td>n = 11</td>
<td>18%</td>
<td>p = 0.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BM, bone marker; n, number of patients; NR, not reported; PMW, post-menopausal women.
Assessment of test accuracy

Most of the evidence available evaluating the accuracy of bone turnover markers was in the form of correlations between changes in bone turnover marker levels and BMD measured using DXA. The studies were extremely heterogeneous precluding the use of any meta-analytical models; the data were therefore presented in tables for each type of treatment (bisphosphonates, teriparatide, raloxifene, strontium ranelate and denosumab) with a brief narrative. Even if the data were not heterogeneous, the usefulness of these correlation data to inform the accuracy of bone turnover marker tests for identifying patients who remain at risk of fracture is limited; this is discussed in more detail in Limitations of the available evidence.

Twenty-seven studies reported the results of correlation analyses, eight reported the results of multiple regression analyses, and four reported both. The \( r \)-values reported were derived using Pearson’s correlation in seven studies and Spearman’s rank correlation in 11; the method was not reported in nine studies. Five studies reported predictive accuracy results in terms of sensitivity, results of receiver operating characteristic (ROC) analyses, or reported sufficient data to produce 2 × 2 tables of test performance. One uncontrolled cohort study reported the difference in bone turnover marker measurements between those with and without a fracture, and another between those who did and did not have a response in BMD to treatment.

### Bisphosphonates

Eighteen studies treated patients using bisphosphonates. The variables that have been correlated, the time points at which they were measured and the patient population receiving treatment are given in Table 4 for each bisphosphonate, along with the results of the correlation analyses. As can be seen from the table, no two studies assessed the same combination of variables and patient population, either within each drug or across bisphosphonates as a whole. Of the 54 \( r \)-values reported for correlations between changes in bone turnover marker and BMD, 19 were statistically significant \((p < 0.05)\); however, all of the correlations were weak \((r < 0.5)\). There were insufficient data for any combination of bone turnover marker, DXA site, time points at which tests were conducted, and patient population to identify any patterns in the data. One study reported correlations between changes in bone turnover markers and fracture incidence; there was a significant treatment by time interaction after the third administration of intravenous zoledronate with sCTX, but no significant association with sP1NP.

Limited data from studies conducting multiple regression analyses indicated that there may be a significant association between the changes in bone turnover markers and either BMD or the incidence of hip or vertebral fractures (Table 5). However, although the studies adjusted for confounding factors, only one adjusted for all of the confounders considered important by the review authors. Predictive ability was assessed using alternative methods in some studies (Table 6); these data provided little evidence regarding the predictive ability of the tests being evaluated.

<table>
<thead>
<tr>
<th>OPPS domain</th>
<th>BM feedback</th>
<th>No BM feedback</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite score</td>
<td>74.3 (14.9)</td>
<td>70.8 (16.0)</td>
<td>3.5 (0.94 to 6.06)</td>
</tr>
<tr>
<td>Confidence score</td>
<td>71.2 (20.7)</td>
<td>70.0 (20.0)</td>
<td>1.2 (–2.16 to 4.56)</td>
</tr>
<tr>
<td>Feel informed</td>
<td>67.8 (21.5)</td>
<td>61.2 (24.3)</td>
<td>6.6 (2.81 to 10.39)</td>
</tr>
<tr>
<td>Motivation score</td>
<td>78.3 (19.4)</td>
<td>76.5 (20.9)</td>
<td>1.8 (–1.53 to 5.13)</td>
</tr>
<tr>
<td>Satisfaction score</td>
<td>79.8 (17.5)</td>
<td>75.5 (19.2)</td>
<td>4.3 (1.26 to 7.34)</td>
</tr>
</tbody>
</table>

BM, bone marker; SD, standard deviation.

### TABLE 4

Score from the OPPS questionnaire as reported in Kung et al. (2009)

<table>
<thead>
<tr>
<th>OPPS domain</th>
<th>Mean (SD)</th>
<th>n</th>
<th>Mean (SD)</th>
<th>n</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite score</td>
<td>74.3 (14.9)</td>
<td>283</td>
<td>70.8 (16.0)</td>
<td>277</td>
<td>3.5 (0.94 to 6.06)</td>
</tr>
<tr>
<td>Confidence score</td>
<td>71.2 (20.7)</td>
<td>285</td>
<td>70.0 (20.0)</td>
<td>277</td>
<td>1.2 (–2.16 to 4.56)</td>
</tr>
<tr>
<td>Feel informed</td>
<td>67.8 (21.5)</td>
<td>287</td>
<td>61.2 (24.3)</td>
<td>278</td>
<td>6.6 (2.81 to 10.39)</td>
</tr>
<tr>
<td>Motivation score</td>
<td>78.3 (19.4)</td>
<td>287</td>
<td>76.5 (20.9)</td>
<td>278</td>
<td>1.8 (–1.53 to 5.13)</td>
</tr>
<tr>
<td>Satisfaction score</td>
<td>79.8 (17.5)</td>
<td>285</td>
<td>75.5 (19.2)</td>
<td>278</td>
<td>4.3 (1.26 to 7.34)</td>
</tr>
</tbody>
</table>

BM, bone marker; SD, standard deviation.
### TABLE 4 Results from the studies reporting correlations between changes in bone turnover marker tests and either changes in DXA, vertebral strength index, or the incidence of fracture in patients being treated with bisphosphonates

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alendronate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>sBALP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)¹⁴</td>
<td>GCS induced 1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>58</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Absolute change FN BMD</td>
<td>63</td>
<td></td>
<td>–0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolute change LS BMD</td>
<td></td>
<td></td>
<td>–0.19</td>
<td></td>
</tr>
<tr>
<td>**Ishijima (2009)⁶⁴</td>
<td>Heterogeneous 6</td>
<td>Per cent change LS BMD</td>
<td>6</td>
<td>45</td>
<td>–0.185*</td>
<td>(–0.457 to 0.119)</td>
</tr>
<tr>
<td>**Watts (2001)³⁵⁵</td>
<td>PMW ≥ 45 years</td>
<td>Per cent change FN BMD</td>
<td>12</td>
<td>180</td>
<td>–0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td>–0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td></td>
<td>–0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>180</td>
<td>–0.36⁶</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td>–0.2⁴</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td></td>
<td>–0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per cent change TB BMD</td>
<td>12</td>
<td>180</td>
<td>–0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td>–0.2⁵</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td></td>
<td>–0.2³</td>
<td></td>
</tr>
<tr>
<td>**Kyd (1998)³⁶⁷</td>
<td>Women 6</td>
<td>Per cent change FN BMD</td>
<td>12</td>
<td>35</td>
<td>–0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>–0.2₅</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per cent change LS BMD</td>
<td></td>
<td></td>
<td>–0.2⁴</td>
<td></td>
</tr>
<tr>
<td>**Reyes-Garcia (2010)³⁸⁸</td>
<td>PMW No statistically signiificant correlations between sBALP and BMD – unclear which time points and DXA sites were analysed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>sP1NP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)¹⁴</td>
<td>GCS induced 1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>71</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td></td>
<td>–0.3⁰</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>70</td>
<td>–0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td></td>
<td>–0.3⁴</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 4  Results from the studies reporting correlations between changes in bone turnover marker tests and either changes in DXA, vertebral strength index, or the incidence of fracture in patients being treated with bisphosphonates (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sCTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)</td>
<td>GCS induced</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>60</td>
<td>–0.32&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>59</td>
<td></td>
<td>–0.29&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>62</td>
<td>–0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>62</td>
<td></td>
<td>–0.28&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Reyes-Garcia (2010)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PMW</td>
<td>3</td>
<td>Change LS BMD</td>
<td>12</td>
<td>46</td>
<td>–0.304&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
<tr>
<td>Kyd (1998)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PMW</td>
<td>52–82 years</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>30</td>
<td>–0.05</td>
</tr>
<tr>
<td>Kyd (1998)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PMW</td>
<td>52–82 years</td>
<td>Per cent change FN BMD</td>
<td>30</td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td><strong>uNTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim (2005)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>PMW</td>
<td>3</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>NR</td>
<td>–0.244</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td>–0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td></td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td>–0.376&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
<tr>
<td>Kyd (1998)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PMW</td>
<td>52–82 years</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>30</td>
<td>–0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>30</td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>Ishijima (2009)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Heterogeneous</td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>6</td>
<td>45</td>
<td>–0.332&lt;sup&gt;a&lt;/sup&gt; &lt;br&gt;(–0.575 to –0.035)</td>
</tr>
<tr>
<td>lwamoto (2005)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>PMW</td>
<td>54–88 years</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>105</td>
<td>–0.20&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>105</td>
<td></td>
<td>–0.341&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>105</td>
<td>–0.338&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>lwamoto (2004)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>PMW</td>
<td>55–88 years</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>85</td>
<td>–0.321&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Alendronate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imai (2009)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>PMW</td>
<td>49–85 years</td>
<td>Vertebral strength index</td>
<td>3</td>
<td>33</td>
<td>0.295</td>
</tr>
<tr>
<td>Masaryk (2002)&lt;sup&gt;i&lt;/sup&gt;</td>
<td>PMW</td>
<td>3</td>
<td>Change LS BMD</td>
<td>12</td>
<td>≥42</td>
<td>–0.310&lt;sup&gt;c&lt;/sup&gt; &lt;br&gt;–0.306&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Change FN BMD</td>
<td></td>
<td></td>
<td>–0.285</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Change TB BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4 Results from the studies reporting correlations between changes in bone turnover marker tests and either changes in DXA, vertebral strength index, or the incidence of fracture in patients being treated with bisphosphonates (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alendronate and risedronate groups reported combined</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armstrong (2007)</td>
<td>Heterogeneous</td>
<td>Unclear</td>
<td>Absolute change spine BMD</td>
<td>20 to 29 months after BM</td>
<td>46</td>
<td>0.25</td>
</tr>
<tr>
<td>Dobnig (2006)</td>
<td>PMW ≥ 60 years</td>
<td>2</td>
<td>Per cent change FN BMD</td>
<td>12</td>
<td>37</td>
<td>−0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td>37</td>
<td>−0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td>37</td>
<td>−0.23</td>
</tr>
<tr>
<td><strong>Ibandronate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hochberg (2010)</td>
<td>PMW 55–80 years</td>
<td>3</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>NR</td>
<td>−0.19b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td>−0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td></td>
<td></td>
<td>−0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>NR</td>
<td>−0.22c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td>−0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td></td>
<td></td>
<td>−0.10</td>
</tr>
<tr>
<td><strong>Zolendronate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delmas (2009)</td>
<td>PMW</td>
<td>12</td>
<td>Fracture incidence</td>
<td>36</td>
<td>No significant association when comparing deciles of sP1NP levels</td>
<td>553</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>Result</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sCTX</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delmas (2009)</td>
<td>PMW</td>
<td>Significant treatment by time interaction after third zolendronate injection</td>
<td>174</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; GCS, glucocorticoid steroid; LS, lumbar spine; n, number of patients; NR, not reported; PMW, post-menopausal women; r, Spearman’s or Pearson’s correlation coefficient; TB, total body.

a Authors reported the use of a simple regression analysis, with the results reported as r-values.
b p < 0.001.
c p < 0.05.
d p < 0.01.
TABLE 5 Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with bisphosphonates

<table>
<thead>
<tr>
<th>Study</th>
<th>Population; treatment</th>
<th>BM</th>
<th>Time point (months)</th>
<th>Dependent variable</th>
<th>Time point (months)</th>
<th>Other independent variables adjusted for</th>
<th>n</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bauer (2004)^39</td>
<td>55–80 years; alendronate, 5 mg/day increased to 10 mg/day</td>
<td>sBALP; 1 SD decrease</td>
<td>12</td>
<td>Incidence vertebral fracture</td>
<td>Final follow-up</td>
<td>Age</td>
<td>1764</td>
<td>OR 0.79 (95% CI 0.66 to 0.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Incidence hip fracture</td>
<td></td>
<td></td>
<td></td>
<td>RH 0.58 (95% CI 0.42 to 0.79)</td>
</tr>
<tr>
<td>Iwamoto (2004)^55</td>
<td>≥ 45, prior spinal fracture; alendronate, 5 mg/day</td>
<td>uNTX</td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>Age, BMI, body weight, baseline BMD, height, other BM results, prior fracture, years since menopause</td>
<td>85</td>
<td>Regression co-efficient: -0.605, R² 0.103</td>
</tr>
<tr>
<td>^Hochberg (2010)^61</td>
<td>55–80 years; ibandronate, 150 mg/month</td>
<td>sCTX</td>
<td>3</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>None</td>
<td>At least 276</td>
<td>R² 0.53^5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change TH BMD</td>
<td></td>
<td></td>
<td></td>
<td>R² 0.46^5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td></td>
<td>R² 0.24^5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change BMD (across all sites)</td>
<td></td>
<td></td>
<td></td>
<td>R² 0.25 to 0.58^5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>Age, baseline BMD</td>
<td>At least 276</td>
<td>R² 0.61^5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change TH BMD</td>
<td></td>
<td></td>
<td></td>
<td>R² 0.58^4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td></td>
<td></td>
<td></td>
<td>R² NR, p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change BMD (across all sites; ≥ 3% a response)</td>
<td>12</td>
<td></td>
<td></td>
<td>R² NR, p &gt; 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per cent change LS BMD</td>
<td></td>
<td></td>
<td></td>
<td>R² 0.60^5</td>
</tr>
<tr>
<td>Study</td>
<td>Population; treatment</td>
<td>BM</td>
<td>Time point (months)</td>
<td>Dependent variable</td>
<td>Time point (months)</td>
<td>Other independent variables adjusted for</td>
<td>n</td>
<td>Result</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------------------</td>
<td>--------</td>
<td>---------------------</td>
<td>--------------------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
<td>-----</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Eastell (2003)(^1)(^3)</td>
<td>Prior spinal fracture; risedronate, 5 mg/day</td>
<td>uCTX</td>
<td>Mean 3–6</td>
<td>Incidence vertebral fracture</td>
<td>12 and 36</td>
<td>Age, baseline BMD, prior fracture</td>
<td>355</td>
<td>p &lt; 0.001 for each time point</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>358</td>
<td>p &lt; 0.05 for each time point (post hoc reanalysis)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>358</td>
<td>p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>uNTX</td>
<td></td>
<td>Incidence vertebral fracture</td>
<td>36</td>
<td></td>
<td>355</td>
<td>p = 0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Incidence non-vertebral fracture</td>
<td></td>
<td></td>
<td>350</td>
<td>p &lt; 0.001 for each time point</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>358</td>
<td>p &lt; 0.05 for each time point (post hoc reanalysis)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350</td>
<td>p &lt; 0.05</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; GCS, glucocorticoid steroid; LS, lumbar spine; n, number of patients; NR, not reported; OR, odds ratio; R\(^2\), regression coefficient; RH, relative hazard; TH, total hip.

a \( p < 0.01 \).
b Hochberg (2004)\(^1\)\(^3\) did not report the regression coefficient, only the \( R^2 \) value and the \( p \)-value associated with the regression coefficient.
c \( p < 0.001 \).
d \( p < 0.05 \).
### TABLE 6 Results from studies using various methods to measure the predictive ability of bone turnover markers to predict changes in BMD or fracture risk in patients being treated with bisphosphonates

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Treatment</th>
<th>Outcome measure; BM time point</th>
<th>Vertebral fracture</th>
<th>No vertebral fracture</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shiraki (2011)</td>
<td>PMW</td>
<td>Alendronate 5 mg/day or 35 mg/week, or rismidronate 2.5 mg/day or 17.5 mg/week</td>
<td>Per cent change in sBALP; mean 4.3 ± 1.9 years where no fracture; mean 3.7 ± 2 years where fracture</td>
<td>−20.3 (61.4) 49</td>
<td>−31.8 (32.0) 153</td>
<td>11.5 (−6.42 to 29.42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change in sBALP; mean 4.3 ± 1.9 years where no fracture; mean 3.7 ± 2 years where fracture</td>
<td>10.8 (17.2) 49</td>
<td>13.4 (15.8) 153</td>
<td>−2.6 (−8.03 to 2.83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change in uNTX; mean 4.3 ± 1.9 years where no fracture; mean 3.7 ± 2 years where fracture</td>
<td>−37.7 (48.2) 58</td>
<td>−31.9 (59.1) 167</td>
<td>−5.8 (−21.10 to 9.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change in uNTX; mean 4.3 ± 1.9 years where no fracture; mean 3.7 ± 2 years where fracture</td>
<td>28.0 (36.3) 58</td>
<td>25.7 (42.4) 167</td>
<td>2.3 (−9.04 to 13.64)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Treatment</th>
<th>BM; BM time point; reference standard</th>
<th>True-positive</th>
<th>False-positive</th>
<th>False-negative</th>
<th>True-negative</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hochberg (2010)</td>
<td>PMW</td>
<td>Ibandronate, 150 mg/monthly</td>
<td>sCTX (cut-off 67%); 3 months; gain LS BMD</td>
<td>136</td>
<td>50</td>
<td>111</td>
<td>104</td>
<td>55.1%</td>
<td>67.5%</td>
</tr>
<tr>
<td></td>
<td>55–80 years</td>
<td></td>
<td>sCTX (cut-off 5%); 3 months; gain LS BMD</td>
<td>245</td>
<td>133</td>
<td>2</td>
<td>21</td>
<td>99.2%</td>
<td>13.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Treatment</th>
<th>BM; time point; reference standard</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitatani (2003)</td>
<td>PMW</td>
<td>Cyclic etidronate; 200 mg/day 2 weeks; 10 weeks off</td>
<td>sBALP; 3 months; gain LS BMD</td>
<td>69</td>
<td>88</td>
<td>0.63</td>
<td>At least 59 patients in analysis (4 dropped out by 24 weeks; analysis 12-week data)</td>
</tr>
</tbody>
</table>

AUC, area under the curve; BM, bone turnover marker; LS, lumbar spine; n, number of patients; NR, not reported; PMW, post-menopausal women; SD, standard deviation.
Overall, the type of data and quality of the evidence base are insufficient to draw any strong conclusion regarding the predictive accuracy of bone turnover marker tests in a population being treated with bisphosphonates.

**Teriparatide**

Ten studies treated patients using teriparatide.\textsuperscript{14,38,42,106,140,146,150,152,159,162} Seven studies administered 20 µg/day (Table 7),\textsuperscript{14,38,42,140,146,150,162} two administered 40 µg/day (Table 8),\textsuperscript{140,159} two reported results for a combined population that received either 20 or 40 µg/day,\textsuperscript{140,152} and one study did not report the dose administered (Table 9).\textsuperscript{106}

Of the nine studies that reported the results of correlation analyses, the $r$-values reported were derived using Pearson’s correlation in one study\textsuperscript{152} and Spearman’s rank correlation in seven,\textsuperscript{14,38,42,106,140,146,159} the method was not reported in one study.\textsuperscript{150} Across 71 reported $r$-values for correlations between changes in bone turnover and BMD in patients treated with 20 µg/day teriparatide, 22 were statistically significant (see Table 7); however, all of them were weak ($r < 0.5$). In patients being treated with 40 µg/day, 9 of 21 reported $r$-values indicated statistically significant, but weak, correlations (see Table 8). When data for patients receiving 20 and 40 µg/day were combined, 12 of 20 reported $r$-values were statistically significant (see Table 9), but, again, all correlations were weak. One of these studies analysed data for the 20 µg/day and 40 µg/day arms separately and combined;\textsuperscript{140} the results for the patients receiving 40 µg/day teriparatide were similar to those of the combined arm, whereas the correlations between changes in bone turnover and BMD in those receiving 20 µg/day were much smaller.\textsuperscript{140} The study that did not report the dose of teriparatide used reported a single non-significant $r$-value (see Table 9).\textsuperscript{106} As with the results for those treated with bisphosphonates, no two studies assessed the same combination of variables and patient population, and there were insufficient data for each combination of bone turnover marker, DXA site, time points at which the tests were conducted, and patient population to identify any patterns in the data.

One study reported correlations between changes in bone turnover markers and dynamic bone parameters determined by bone biopsy in patients receiving 20 µg/day (see Table 7).\textsuperscript{150} This study suggests that sCTX and sP1NP may be positively correlated with improvements in dynamic bone parameters measured after 24 months of treatment, particularly when the bone turnover marker was also measured at 24 months (no results were presented for the associations with 3, 6 and 12 months’ bone turnover marker tests; however, the study was reported as an abstract and there was no response from the study authors to a request for further information); moderate to strong correlations were observed between the 24-month change in sCTX and sP1NP and the 24-month change in activation frequency ($r = 0.69$ and $r = 0.73$, respectively).\textsuperscript{150} A second study reported correlations between changes in bone turnover markers and bone structural and dynamic parameters for a population that combined patients receiving 20 or 40 µg/day (see Table 9).\textsuperscript{152} This study reported significant correlations between the changes in sBALP conducted at 1 month and some structural parameters, with a strong correlation between the change in sBALP and mean wall thickness ($r = 0.73$); there were no significant correlations with dynamic parameters. There were no significant correlations between structural or dynamic parameters and uNTX.\textsuperscript{152} It is possible that significant correlations could have been detected if the bone turnover marker test had been delayed beyond 1 month, as demonstrated in the first study,\textsuperscript{150} or if the analyses included a greater number of patients. In addition, as demonstrated by Chen et al.,\textsuperscript{140} the strength of the correlation will be affected by the combining of data for 20 and 40 µg/day, particularly in such a small study.
TABLE 7 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD, bone biopsy results, or the incidence of fracture in patients being treated with teriparatide 20µg/day

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sBALP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)14</td>
<td>GCS induced</td>
<td>1</td>
<td>Absolute change FN BMD 18</td>
<td>52</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD</td>
<td>52</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change FN BMD 58</td>
<td>59</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD 59</td>
<td></td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Chen (2005)140</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change FN BMD 12</td>
<td>148</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD 18</td>
<td>132</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change FN BMD 12</td>
<td>148</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD 18</td>
<td>132</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change FN BMD 12</td>
<td>148</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD 18</td>
<td>132</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change FN BMD 12</td>
<td>148</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD 18</td>
<td>132</td>
<td>−0.05</td>
<td></td>
</tr>
<tr>
<td>Tsujimoto (2011)46</td>
<td>≥ 55 years</td>
<td>1</td>
<td>Per cent change LS BMD 12</td>
<td>121</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>−0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Per cent change LS BMD 12</td>
<td>121</td>
<td>−0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>−0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>−0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD 12</td>
<td>121</td>
<td>−0.20*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>−0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>−0.23*</td>
<td></td>
</tr>
<tr>
<td><strong>sCTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)14</td>
<td>GCS induced</td>
<td>1</td>
<td>Absolute change LS BMD 18</td>
<td>57</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD 57</td>
<td>57</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD 59</td>
<td>59</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD 58</td>
<td>58</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Stepan (2008)55</td>
<td>PMW</td>
<td>1</td>
<td>Per cent change double-labelled perimeter 24</td>
<td>35</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change MS/BS 35</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change AcF 35</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change BFR 35</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>Per cent change double-labelled perimeter 35</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change MS/BS 35</td>
<td>0.45*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change AcF 35</td>
<td>0.69*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change BFR 35</td>
<td>0.44*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 7 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD, bone biopsy results, or the incidence of fracture in patients being treated with teriparatide 20µg/day

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsujimoto (2011)</td>
<td>≥ 55 years</td>
<td>1</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>121</td>
<td>–0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>120</td>
<td>–0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>120</td>
<td>–0.23*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Per cent change LS BMD</td>
<td>121</td>
<td>121</td>
<td>–0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>120</td>
<td>–0.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>120</td>
<td>–0.26*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>121</td>
<td>121</td>
<td>–0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td>120</td>
<td>–0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change hip BMD</td>
<td>120</td>
<td>120</td>
<td>–0.24*</td>
</tr>
<tr>
<td>sP1NP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blumsohn (2011)</td>
<td>PMW ≥ 55 years</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>24</td>
<td>414</td>
<td>0.213*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change hip BMD</td>
<td>401</td>
<td>401</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>414</td>
<td>414</td>
<td>0.365*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change hip BMD</td>
<td>401</td>
<td>401</td>
<td>0.141*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD</td>
<td>414</td>
<td>414</td>
<td>0.117*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>401</td>
<td>401</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change hip BMD</td>
<td>401</td>
<td>401</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change LS BMD</td>
<td>414</td>
<td>414</td>
<td>0.219*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>401</td>
<td>401</td>
<td>0.111*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change hip BMD</td>
<td>401</td>
<td>401</td>
<td>0.107*</td>
</tr>
<tr>
<td>sP1NP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burshell (2010)</td>
<td>GCS induced</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>77</td>
<td>0.33*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>77</td>
<td>77</td>
<td>0.34*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>77</td>
<td>77</td>
<td>0.23*</td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>132</td>
<td>0.26*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>148</td>
<td>148</td>
<td>–0.04</td>
</tr>
<tr>
<td>Miller (2008)</td>
<td>PMW 51–85 years;</td>
<td>3</td>
<td>Areal BMD</td>
<td>12</td>
<td>NR</td>
<td>NRd</td>
</tr>
<tr>
<td></td>
<td>prior BP treatment</td>
<td></td>
<td>Volumetric BMD of spine and hip</td>
<td>NR</td>
<td>NR</td>
<td>0.45*</td>
</tr>
</tbody>
</table>

* Continued...
TABLE 7 Results from studies reporting correlations between changes in bone turnover markers and changes in BMD, bone biopsy results, or the incidence of fracture in patients being treated with teriparatide 20µg/day (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stepán (2008)</td>
<td>PMW</td>
<td>1</td>
<td>Per cent change double-labelled perimeter</td>
<td>24</td>
<td>35</td>
<td>0.39&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change MS/BS</td>
<td>35</td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change AcF</td>
<td>35</td>
<td></td>
<td>0.49&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change BFR</td>
<td>35</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>Per cent change double-labelled perimeter</td>
<td>35</td>
<td></td>
<td>0.39&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change MS/BS</td>
<td>35</td>
<td></td>
<td>0.48&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change AcF</td>
<td>35</td>
<td></td>
<td>0.73&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change BFR</td>
<td>35</td>
<td></td>
<td>0.47&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tsujimoto (2011)</td>
<td>≥ 55 years</td>
<td>1</td>
<td>Per cent change LS BMD</td>
<td>12</td>
<td>121</td>
<td>0.56&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td></td>
<td>−0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change Hip BMD</td>
<td>120</td>
<td></td>
<td>0.21&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Per cent change LS BMD</td>
<td>121</td>
<td></td>
<td>0.36&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td></td>
<td>−0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change Hip BMD</td>
<td>120</td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Per cent change LS BMD</td>
<td>121</td>
<td></td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change FN BMD</td>
<td>120</td>
<td></td>
<td>−0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Per cent change Hip BMD</td>
<td>120</td>
<td></td>
<td>−0.18</td>
</tr>
<tr>
<td>uNTX</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>132</td>
<td>−0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>148</td>
<td>−0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>132</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>148</td>
<td>−0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>132</td>
<td>−0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>148</td>
<td>−0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>132</td>
<td>−0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>148</td>
<td>−0.10</td>
</tr>
</tbody>
</table>

AcF, activation frequency; BFR, bone formation rate; BM, bone turnover marker; FN, femoral neck; GCS, glucocorticoid steroid; LS, lumbar spine; MS/BS, mineralising surface; n, number of patients; NR, not reported; PMW, post-menopausal women; r, Spearman’s or Pearson’s correlation coefficient; TB, total body.

<sup>a</sup> p < 0.05.
<sup>b</sup> p < 0.01.
<sup>c</sup> p < 0.001.
<sup>d</sup> No significant correlation (data not reported).
TABLE 8  Results from studies reporting correlations between changes in bone turnover markers and changes in BMD in patients being treated with teriparatide 40µg/day

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sBALP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.26*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.17</td>
</tr>
<tr>
<td>Lane (2000)</td>
<td>PMW GCS induced</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>12</td>
<td>28</td>
<td>0.39*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>12</td>
<td>28</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>12</td>
<td>28</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Associations were comparable when the analyses were performed for 24-month changes in spine BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sP1NP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.38*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.24*</td>
</tr>
<tr>
<td>uNTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.30*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.32*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.12</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; LS, lumbar spine; n, number of patients; PMW, post-menopausal women; r, Spearman’s or Pearson’s correlation coefficient.

a. p < 0.05.
<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>20 or 40 µg/day</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>sBALP</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.19*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.22*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.17*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.24*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.14*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.12*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.14*</td>
</tr>
<tr>
<td>Dobnig (2005)</td>
<td>PMW</td>
<td>1</td>
<td>2D cortical thickness</td>
<td>22</td>
<td>16</td>
<td>−0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2D marrow star volume</td>
<td>15</td>
<td>15</td>
<td>−0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2D mean wall thickness</td>
<td>17</td>
<td>17</td>
<td>0.73*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2D trabecular bone volume</td>
<td>16</td>
<td>16</td>
<td>0.58*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3D CD</td>
<td>19</td>
<td>19</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3D cortical thickness</td>
<td>15</td>
<td>15</td>
<td>−0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3D SMI</td>
<td>19</td>
<td>19</td>
<td>−0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3D trabecular bone volume</td>
<td>19</td>
<td>19</td>
<td>0.54*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3D trabecular number</td>
<td>19</td>
<td>19</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>trabecular thickness</td>
<td>19</td>
<td>19</td>
<td>0.49*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2D dynamic parameters</td>
<td>NR</td>
<td>NR</td>
<td>NR*</td>
</tr>
<tr>
<td><strong>sP1NP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>127</td>
<td>0.40*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>137</td>
<td>0.15*</td>
</tr>
<tr>
<td><strong>uNTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>PMW</td>
<td>1</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.19*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.17*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>Absolute change LS BMD</td>
<td>18</td>
<td>259</td>
<td>0.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>12</td>
<td>285</td>
<td>0.06</td>
</tr>
<tr>
<td>Dobnig (2005)</td>
<td>PMW</td>
<td>1</td>
<td>Structural parameters</td>
<td>22</td>
<td>NR</td>
<td>NR*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dynamic parameters</td>
<td>NR</td>
<td>NR</td>
<td>NR*</td>
</tr>
</tbody>
</table>
No studies were identified that reported the results of multiple regression analyses in patients being treated with teriparatide. Predictive ability was assessed using alternative methods in some studies (Table 10); these data provide little evidence regarding the predictive ability of the tests being evaluated.

Overall, the quality of the evidence base is insufficient to draw any strong conclusion regarding the predictive accuracy of bone turnover marker tests in a population being treated with teriparatide. Further, the lack of evidence of a correlation between changes in bone turnover markers and BMD does not indicate a lack of association between changes of bone turnover and fracture risk.

**Raloxifene**

Three studies treated patients using raloxifene. Of the 12 reported \( r \)-values, none was statistically significant and all correlations were weak (Table 11). There is some evidence that changes in sBALP may be associated with the incidence of fracture (Table 12); however, this was assessed in a single study that did not distinguish between women receiving 60 mg/day and 120 mg/day of raloxifene; it is unclear whether or not the significant association seen in this study of raloxifene for the combined dose would be evident with both doses if analyses had been conducted separately for each dose. The evidence base is insufficient to draw conclusions regarding the predictive accuracy of bone turnover marker tests in a population being treated with raloxifene.

**Strontium ranelate**

Two studies evaluated patients treated using strontium ranelate. Of the seven reported \( r \)-values, six were statistically significant (Table 13); sP1NP and sCTX showed moderate correlations with absolute changes in BMD (\( r = 0.615 \) and 0.56, respectively). A single study reported the results of a multiple regression analysis and correlation (Table 14). Although the correlation between changes in sBALP and changes in BMD at the femoral neck were found to be weak and non-significant, when evaluated in the multiple regression analysis on a percentage change basis a significant association was evident; several explanatory variables were included in the multiple regression analysis. The opposite was true of sCTX, where significant but weak correlations were not evident in the subsequent multiple regression analysis.

Predictive ability was assessed using area under the curve (AUC) in one study (Table 15); these data provide little evidence regarding the predictive ability of the tests being evaluated. Overall, the evidence base is insufficient to draw conclusions regarding the predictive accuracy of bone turnover marker tests in a population being treated with strontium ranelate.

---

**TABLE 9** Results from studies reporting correlations between changes in bone turnover markers and changes in BMD or bone biopsy results in patients being treated with 20 or 40µg/day teriparatide, or where the dose was unclear (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>( n )</th>
<th>( r )-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S</em>P1NP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddiqi (2010)</td>
<td>Unclear</td>
<td>1</td>
<td>Per cent change LS BMD</td>
<td>18</td>
<td>28</td>
<td>0.093</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; CD, connectivity density; FN, femoral neck; GCS, glucocorticoid steroid; LS, lumbar spine; MS/BS, mineralising surface; \( n \), number of patients; NR, not reported; PMW, post-menopausal women; \( r \), Spearman’s or Pearson’s correlation coefficient; SMI, structural model index; TB, total body.

\( a \) \( p < 0.05 \).

\( b \) \( p < 0.001 \).

\( c \) No significant correlation – data not reported.
<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Dose</th>
<th>BM: time point; reference standard</th>
<th>AUC</th>
<th>Study</th>
<th>Population</th>
<th>Dose</th>
<th>BM: time point; reference standard</th>
<th>Mean change in uNTX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane</td>
<td>PMW, GCS induced, n = 28</td>
<td>40 µg/day</td>
<td>sBALP: 1 month; 12-month gain spinal BMD</td>
<td>0.773</td>
<td>Chen</td>
<td>PMW, n = unclear</td>
<td>20 or 40 µg/day</td>
<td>sBALP: 1 month; 18-month gain LS BMD</td>
<td>Sensitivity: 35%</td>
</tr>
<tr>
<td>(2000)59</td>
<td></td>
<td></td>
<td>sBALP: 3 months; 18-month gain LS BMD</td>
<td>0.704</td>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 18-month gain LS BMD</td>
<td>Sensitivity: 44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 12-month gain spinal BMD</td>
<td>0.812</td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 1 month; 18-month gain LS BMD</td>
<td>Sensitivity: 7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 1 month; 12-month gain spinal BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sBALP: 3 months; 18-month gain LS BMD</td>
<td>Sensitivity: 44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 12-month gain spinal BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sP1NP: 3 months; 18-month gain LS BMD</td>
<td>Sensitivity: 69%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 1 month; 12-month gain spinal BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 1 month; 18-month gain LS BMD</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 3 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 3 months; 18-month gain LS BMD</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 6 months; 18-month gain LS BMD</td>
<td>22%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 1 month; 12-month gain spinal BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 3 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 1 month; 12-month gain spinal BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 3 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 6 months; 18-month gain LS BMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uNTX: 12 months; 18-month gain LS BMD</td>
<td>22%</td>
</tr>
</tbody>
</table>

AUC, area under the curve; BM, bone turnover marker; Cr, creatinine; GCS, glucocorticoid steroid; LS, lumbar spine; n, number of patients; PMW, post-menopausal women.
### Table 11: Results from studies reporting Pearson’s correlation coefficients for analyses associating changes in bone turnover markers and changes in BMD in patients being treated with raloxifene

<table>
<thead>
<tr>
<th>BM Marker</th>
<th>Patient Population</th>
<th>BM Time Point (months)</th>
<th>Correlated with BMD Change (months)</th>
<th>Comparator Time Point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>sBALP</td>
<td>Majima (2008)</td>
<td>3</td>
<td>Absolute change LS BMD 6</td>
<td>63</td>
<td>63</td>
<td>-0.233</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>63</td>
<td>63</td>
<td>-0.186</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD 12</td>
<td>63</td>
<td>63</td>
<td>-0.159</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>63</td>
<td>63</td>
<td>-0.156</td>
</tr>
<tr>
<td>sNTX</td>
<td>Majima (2008)</td>
<td>3</td>
<td>Absolute change LS BMD 6</td>
<td>63</td>
<td>63</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>63</td>
<td>63</td>
<td>0.201</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>Absolute change LS BMD 12</td>
<td>63</td>
<td>63</td>
<td>-0.237</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td>63</td>
<td>63</td>
<td>0.097</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; LS, lumbar spine; n, number of patients; PMW, post-menopausal women; r, Spearman’s or Pearson’s correlation coefficient.
### TABLE 12  Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with raloxifene

<table>
<thead>
<tr>
<th>Study</th>
<th>Population; dose</th>
<th>BM</th>
<th>Time point (months)</th>
<th>Dependent variable</th>
<th>Time point (months)</th>
<th>Other independent variables adjusted for</th>
<th>n</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reginster (2004)</td>
<td>Mixed; 60 mg/day</td>
<td>sBALP</td>
<td>12</td>
<td>Incidence vertebral fracture</td>
<td>36</td>
<td>NR (unclear)</td>
<td>967</td>
<td>Slope 0.0056 (95% CI 0.0003 to 0.0109)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sP1NP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Slope 0.0085 (95% CI 0.0021 to 0.015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sCTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Slope 0.0027 (95% CI −0.0014 to 0.0068)</td>
</tr>
<tr>
<td>Bjarnason (2001)</td>
<td>Mixed; 60 or 120 mg/day</td>
<td>sBALP</td>
<td>6</td>
<td>At least one new vertebral fracture</td>
<td>36</td>
<td>Age, baseline BMD, BMI, prior fracture, smoking status</td>
<td>1534</td>
<td>OR 0.63 (CI 0.5 to 0.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OR 0.75 (CI 0.62 to 0.92)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR fracture significantly greater in patients in upper sBALP tertile than lower at 6 (p = 0.036) and 12 (p = 0.045) months</td>
</tr>
<tr>
<td></td>
<td></td>
<td>uCTX</td>
<td>6</td>
<td>At least one new vertebral fracture</td>
<td>36</td>
<td>Age, baseline BMD, BMI, prior fracture, smoking status</td>
<td>1451</td>
<td>OR 0.91 (CI 0.73 to 1.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OR 0.81 (CI 0.64 to 1.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR fracture not significantly greater in patients in upper sCTX tertile than lower at 6 (p = 0.49) and 12 (p = 0.74) months</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; n, number of patients; NR, not reported; RR, relative risk.
**TABLE 13** Results from studies reporting Spearman’s rank correlations coefficients for analyses associating changes in bone turnover markers and changes in BMD in patients being treated with strontium ranelate

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>n</th>
<th>r-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sBALP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bruyere (2010)</td>
<td>PMW ≥ 50 years</td>
<td>3</td>
<td>Absolute change</td>
<td>36</td>
<td>1737</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FN BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>sP1NP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro-Alvarez (2010)</td>
<td>PMW</td>
<td>Unclear</td>
<td>Absolute change</td>
<td>12 or 24 (unclear)</td>
<td>66</td>
<td>0.615*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>sCTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bruyere (2010)</td>
<td>PMW ≥ 50 years</td>
<td>3</td>
<td>Absolute change</td>
<td>36</td>
<td>1737</td>
<td>–0.09*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change</td>
<td>1737</td>
<td>1737</td>
<td>–0.09*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FN BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro-Alvarez (2010)</td>
<td>PMW</td>
<td>Unclear</td>
<td>Absolute change</td>
<td>12 or 24 (unclear)</td>
<td>66</td>
<td>0.56*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>uNTX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bruyere (2010)</td>
<td>PMW ≥ 50 years</td>
<td>3</td>
<td>Absolute change</td>
<td>36</td>
<td>1737</td>
<td>–0.06*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute change</td>
<td>1737</td>
<td>1737</td>
<td>–0.06*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FN BMD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; LS, lumbar spine; n, number of patients; PMW, post-menopausal women; r, Spearman’s or Pearson’s correlation coefficient.

a  p < 0.05.
b  p < 0.001.

**DOI:** 10.3310/hta18110  **HEALTH TECHNOLOGY ASSESSMENT 2014 VOL. 18 NO. 11**
Table 14: Results from the studies reporting regression analyses in which changes in the level of at least one bone turnover marker of interest was included as an independent variable in patients being treated with strontium ranelate

<table>
<thead>
<tr>
<th>Study</th>
<th>Population; dose</th>
<th>BM</th>
<th>Time point (months)</th>
<th>Dependent variable</th>
<th>Other independent variables adjusted for</th>
<th>n</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruyere (2010)$^{a,b}$</td>
<td>≥ 50 years; 2 g/day</td>
<td>sBALP (1% change)</td>
<td>3</td>
<td>Absolute change LS BMD</td>
<td>36</td>
<td>Age, baseline BMD, other BM results, prior fracture</td>
<td>1737</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td>5.2$^{b}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>sCTX (1% change)</td>
<td></td>
<td>Absolute change LS BMD</td>
<td></td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>uNTX (1% change)</td>
<td></td>
<td>Absolute change LS BMD</td>
<td></td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute change FN BMD</td>
<td></td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>sBALP (change 1 SD)</td>
<td></td>
<td>Vertebral fracture risk reduction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>sCTX (change 1 SD)</td>
<td></td>
<td>8% (95% CI −4% to 20%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>uNTX (change 1 SD)</td>
<td></td>
<td>1% (95% CI −9% to 11%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; LS, lumbar spine; n, number of patients; $R^2$, regression coefficient.

- Although this study reported a column for the regression coefficient, the figures presented did not appear to be $\beta$-coefficients, and therefore only the $R^2$ values and the $p$-values associated with the regression coefficients were extracted.
- $p < 0.001$. 

CLINICAL EFFECTIVENESS AND COST-EFFECTIVENESS EVIDENCE
Denosumab

One study treated patients using denosumab.43 Of the six reported $r$-values, five were statistically significant (Table 16); all correlations were weak ($r < 0.50$). No studies were identified that reported the results of multiple regression analyses or results from any alternative methods of analysis. There is insufficient evidence to draw any conclusions regarding the predictive accuracy of bone turnover marker tests in a population being treated with denosumab.

Assessment of test reliability and reproducibility

Four studies reported S/N ratio for a bone turnover marker in those being treated with either etidronate,156 teriparatide42 or raloxifene.147,164 Each study calculated the S/N ratio differently, making comparisons across studies difficult (Table 17). Within-study comparisons show that sP1NP at 2 weeks had a lower S/N ratio than sCTX at 2 weeks but a higher S/N ratio at 25 weeks,147 and sP1NP had a greater S/N ratio than sBALP when measured at 6 months.42 Blumsohn et al.42 reported the intraclass correlation coefficients measured at two time points between 3 and 14 days apart, which were 0.988 for sBALP and 0.983 for sP1NP.

### TABLE 15

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Dose</th>
<th>BM: time point; reference standard: time point</th>
<th>AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruyere (2010)161</td>
<td>PMW ≥ 50 years, $n = 1737$</td>
<td>2 g/day</td>
<td>sBALP: 3 months; vertebral fracture: 3 years</td>
<td>0.51 (0.47 to 0.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sBALP: 3 months; non-vertebral fracture: 3 years</td>
<td>0.51 (0.47 to 0.57)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sCTX: 3 months; vertebral fracture: 3 years</td>
<td>0.48 (0.43 to 0.53)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sCTX: 3 months; non-vertebral fracture: 3 years</td>
<td>0.46 (0.42 to 0.51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>uNTX: 3 months; vertebral fracture: 3 years</td>
<td>0.52 (0.47 to 0.57)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>uNTX: 3 months; non-vertebral fracture: 3 years</td>
<td>0.47 (0.42 to 0.51)</td>
</tr>
</tbody>
</table>

$n$, number of patients; PMW, post-menopausal women.

### TABLE 16

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient population</th>
<th>BM</th>
<th>BM time point (months)</th>
<th>Correlated with</th>
<th>Comparator time point (months)</th>
<th>$n$</th>
<th>$r$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastell (2011)43</td>
<td>PMW 60–90 years</td>
<td>sBALP</td>
<td>6</td>
<td>Change LS BMD</td>
<td>36</td>
<td>73 or 89 (unclear)</td>
<td>$-0.26^a$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change hip BMD</td>
<td></td>
<td></td>
<td>$-0.06$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sP1NP</td>
<td></td>
<td>Change LS BMD</td>
<td></td>
<td></td>
<td>$-0.42^b$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change hip BMD</td>
<td></td>
<td></td>
<td>$-0.47^b$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sCTX</td>
<td></td>
<td>Change LS BMD</td>
<td></td>
<td></td>
<td>$-0.24^a$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change hip BMD</td>
<td></td>
<td></td>
<td>$-0.44^b$</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; FN, femoral neck; LS, lumbar spine; $n$, number of patients; PMW, post-menopausal women; $r$, Spearman’s or Pearson’s correlation coefficient.

a $p < 0.05$.

b $p < 0.001$. 
Clowes et al.\textsuperscript{147} reported the intraindividual coefficients of variation (CVs), which were 13.2 for sP1NP and 22.9 for sCTX.

### Results of the systematic review of cost-effectiveness

No studies met the inclusion criteria for the systematic review of bone turnover marker monitoring strategies. A list of excluded studies with the reasons for exclusion is given in Appendix 3.

### Discussion

This systematic review set out to (1) determine the clinical effectiveness of monitoring regimens that included at least one bone turnover marker, over and above monitoring regimens where a bone turnover marker was not used, and (2) identify which bone turnover marker, if any, should be introduced into routine practice for the monitoring of response to osteoporosis treatments with the aim of identifying

### TABLE 17 Signal to noise ratios reported for three bone turnover markers across studies with diverse populations and treatment regimens

<table>
<thead>
<tr>
<th>Study</th>
<th>Population; treatment</th>
<th>Calculation used</th>
<th>n</th>
<th>S/N ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bone formation markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sBALP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitatani (2003)\textsuperscript{156}</td>
<td>PMW; 200 mg/day etidronate for 2 weeks; 10 weeks off</td>
<td>12-week change in BM/LSC</td>
<td>Min. 59 (max. 63)</td>
<td>Mean 1.3 (SD 1.2)</td>
</tr>
<tr>
<td>Sarker (2004)\textsuperscript{154}</td>
<td>PMW; raloxifene, 60 or 120 mg/day</td>
<td>(% change BM raloxifene − % change BM placebo)/(SQRT(population variance % changes raloxifene − population variance % changes placebo)); BM measurements at 12 months</td>
<td>2503</td>
<td>0.27 (95% CI 0.19 to 0.36)</td>
</tr>
<tr>
<td>Blumsohn (2011)\textsuperscript{42}</td>
<td>PMW ≥ 55 years; teriparatide, 20 µg/day</td>
<td>‘Signal’ = absolute change in log-transformed values; ‘noise’ = within-subject biological variability; BM measurement at 6 months</td>
<td>83</td>
<td>8</td>
</tr>
<tr>
<td><strong>sP1NP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clowes (2003)\textsuperscript{147}</td>
<td>Unclear; raloxifene, 60 mg/day</td>
<td>Per cent change at 2 weeks/intraindividual CV</td>
<td>22</td>
<td>Mean 0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per cent change at 25 weeks/intraindividual CV</td>
<td>22</td>
<td>Mean 2.7</td>
</tr>
<tr>
<td>Blumsohn (2011)\textsuperscript{42}</td>
<td>PMW ≥ 55 years; teriparatide, 20 µg/day</td>
<td>Absolute change in log-transformed values : within-subject biological variability; BM measurement at 6 months</td>
<td>83</td>
<td>12.4</td>
</tr>
<tr>
<td><strong>Bone resorption markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sCTX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clowes (2003)\textsuperscript{147}</td>
<td>Unclear; raloxifene, 60 mg/day</td>
<td>Per cent change at 2 weeks/intraindividual CV</td>
<td>22</td>
<td>Mean 1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per cent change at 25 weeks/intraindividual CV</td>
<td>22</td>
<td>Mean 1.7</td>
</tr>
</tbody>
</table>

BM, bone turnover marker; CV, coefficient of variation; FN, femoral neck; LS, lumbar spine; LSC, least significant change; max., maximum; min., minimum; n, number of patients; PMW, post-menopausal women; SD, standard deviation; SQRT, square root.

Clowes et al.\textsuperscript{147} reported the intraindividual coefficients of variation (CVs), which were 13.2 for sP1NP and 22.9 for sCTX.
non-responders to treatment and informing treatment choice. In order to address these questions we sought a range of study designs. No studies were identified that evaluated the cost-effectiveness of bone turnover marker monitoring strategies, and therefore the focus of this discussion is the clinical effectiveness evidence.

Randomised controlled trials were included if they compared monitoring with or without a bone turnover marker, the feedback of results of bone turnover tests, or monitoring with different bone turnover markers. We also searched for evidence on the accuracy, reliability and reproducibility of these tests in people with osteoporosis receiving treatment. There are a large number of RCTs that compare different treatment regimens, using bone turnover markers as an outcome to determine treatment response; we excluded these trials. Although such trials of treatment effectiveness indicate the magnitude of change in a bone turnover marker in response to treatment and potentially identify treatment non-responders, without a direct analysis of the association of the change in the bone turnover marker with a measure of fracture risk (such as BMD, incidence of fracture or biopsy), there is no evidence that the bone turnover markers have accurately identified non-responders. We therefore restricted our review of test accuracy to those studies that presented an analysis that directly compared the results of bone turnover markers with a measure of fracture risk.

The review identified a number of studies that used a range of methods to evaluate the clinical effectiveness and predictive value of bone turnover marker tests. Unfortunately, no studies meeting the inclusion criteria investigated the impact of bone turnover marker monitoring on patient management and treatment choices.

**Key findings**

The included RCTs evaluated the impact of the feedback of results of bone turnover marker tests on QoL and/or adherence, compliance and/or persistence with bisphosphonate therapy. Although this is an important area of research, the usefulness of these data is limited because of the treatment regimens administered and the high baseline compliance and persistence rates; these issues are discussed further in *Limitations of the available evidence*, below. There was some evidence that the feedback of bone turnover marker results, and the message given, impacted on persistence with treatment; a positive result seems to encourage persistence and a negative result to discourage it. This may not be the response expected to the feedback of such results in clinical practice. It may be expected that a message that treatment was not working could encourage adherence rather than discourage it, assuming that the potential risk of fracture is highlighted to the patient. It is worth noting that the RCT evaluating this, firstly, was conducted in older women aged 65–80 years, and may not be reflective of the response of younger post-menopausal women or of those with secondary osteoporosis, and, secondly, used daily risedronate which is a regimen that is not commonly used in current clinical practice.

Two of the included RCTs also reported on the QoL using the OPPS questionnaire; these studies reported small improvements for those patients receiving feedback in the overall, feeling informed and satisfaction scores and the confidence score, compared with those who did not receive feedback. Although this is based on only two studies, one received questionnaires from 563 of the 596 participants recruited, and the other study recruited 585 participants although it was not clear how many of these returned questionnaires. A study identified during the scoping review of modelling methods explored the impact of monitoring antiresorptive treatments using a bone turnover marker on quality-adjusted life-years (QALYs). The decision tree (up to 3 months) and Markov model (3 months to 5 years) was based on a 60-year-old woman with post-menopausal osteoporosis with a total hip T-score of –3, no concurrent disease, and second-generation bisphosphonate therapy for 5 years. The comparators were no monitoring beyond a simple short-term follow-up to rule out adverse reactions, and monitoring with a serum bone resorption marker after 3 months of treatment. This modelling study showed small improvements in QALYs when adherence to treatment was assumed to be the same in both monitoring strategies, with the improvement being greater if monitoring improved adherence.
The majority of the evidence for the review of clinical effectiveness was based around the predictive accuracy of bone turnover marker tests in the form of correlations with changes in BMD; the limitations of these data are discussed in Limitations of the available evidence, below. Given the extreme heterogeneity across the studies, it was impossible to determine trends for any particular treatment–bone turnover marker monitoring combination. Studies correlating changes in bone turnover with either biopsy results or fracture outcomes were uncommon, with two studies using biopsy results as the comparator and seven using fracture. The use of biopsy in all patients in a study is unlikely to be considered ethical because of the invasive nature of the procedure and the risk of complications. In addition, there are several limitations to the use of fracture outcomes as the comparator: the outcome is relatively rare, resulting in the need for larger studies; there needs to be a longer duration of follow-up in order to detect the outcome; and attrition is likely to be a problem over the time period required. These limitations may be the reason that DXA is most commonly used for the comparator. Data from the studies that used either biopsy or fracture were heterogeneous in terms of patient population and treatment regimen, and method of data analysis. The results of these studies were inconsistent and there were insufficient data to determine trends for any particular treatment–bone turnover marker monitoring combination.

There was no evidence available to inform the question as to the clinical effectiveness of treatment monitoring including a bone turnover marker over and above monitoring regimens where a bone turnover marker was not used. Further, the evidence available relating to the predictive accuracy, reliability and reproducibility was heterogeneous and of low quality, precluding the ability to draw any conclusions regarding the choice of bone turnover marker for use in monitoring in routine clinical practice.

Strengths and limitations of the review
The systematic review was based on an extensive search with well-designed search strategies. Abstracts were included when there were sufficient data to be extracted; authors of all potentially included abstracts were contacted in an attempt to identify full publications and/or obtain unpublished data. In addition, no study was excluded based on date of conduct or language of publication. Studies were selected using inclusion criteria defined a priori. We included RCTs of any size in the review of the clinical effectiveness of bone turnover marker monitoring. However, given (1) the large number of non-randomised studies and randomised studies from which cohorts were derived, from which data on test accuracy, reliability and reproducibility could have been extracted, (2) the apparent poor quality of this evidence base, and (3) the time constraints on the project, an additional criterion not specified in the protocol was applied. We excluded non-randomised studies and derived cohorts that included fewer than 20 osteoporotic patients, receiving one of the treatments of interest, in all the analyses of outcomes relevant to this review. This cut-off was used to distinguish between a case series and a cohort study, excluding from the review the lowest levels of evidence. This is the only change made to the protocol. Although the studies were too heterogeneous to pool using statistical meta-analytical models, we conducted a narrative synthesis in an attempt to summarise the evidence available.

Limitations of the available evidence
As with all systematic reviews, the reliability and generalisability of the results are governed by the quality and quantity of the evidence available. Although we included 42 studies, all were considered to be at a high or unclear risk of bias, and therefore of low quality. Furthermore, no RCTs were identified that addressed one of the primary aims of the review: to evaluate the clinical effectiveness of monitoring regimens that included at least one bone turnover marker, over and above monitoring regimens where a bone turnover marker was not used. Where RCTs were conducted, these reported primarily on compliance.

Adherence to treatment (in terms of both compliance and persistence), particularly with oral bisphosphonates, is a considerable problem in the management of patients with osteoporosis. Oral bisphosphonates are associated with gastrointestinal adverse events, and require adherence to specific instructions for administration to maximise absorption and bioavailability that many patients find inconvenient. Non-adherence to treatment regimens will result in a non- (or inadequate) response to
treatment and a continued higher risk of fracture.173,174 A recent systematic review of 27 observational studies, most of which were retrospective database analyses, reported a rate of fracture ranging from 6% to 38% in patients who were non-compliant, and 5% to 19% in patients who were non-persistent with osteoporosis medications.175 The meta-analysis conducted included 12 of the studies and indicated a statistically significant increased risk of fracture of approximately 30% in patients who were non-compliant (OR 1.29, 95% CI 1.22 to 1.38), and 30% to 40% in patients who were non-persistent (OR 1.40, 95% CI 1.29 to 1.52) with treatment.175 Most of the patients in the review received bisphosphonates; however, raloxifene, strontium ranelate, calcitonin, HRT, vitamin D and calcium were included in the regimens administered.175

In the five RCTs that met the inclusion criteria for this review,56,133,143,148,149 a very high proportion of patients were adherent to medication; this is unlikely to be representative of clinical practice. Why there was such a high rate of compliance in the trials is unclear. It could be due to the use of a once-monthly dosing regimen used in most of the trials rather than daily or weekly dosing, as less frequent dosing is generally preferred by patients and can result in increased adherence.32,173,176 Other potential reasons are: the short duration of follow-up used in most of the studies; that patients giving consent to be part of the trial are more likely to be compliant; and/or the increased attention and resulting patient awareness that being part of a trial may result in. It is, therefore, unclear how the feedback of bone turnover marker results would impact on the population seen in generally clinical practice, which is likely to be less compliant and/or persistent.

Most of the evidence available relating to the accuracy of the bone turnover markers was in the form of correlations with BMD, measured using DXA. Some of the correlations, but not all, were statistically significant. The fact that a correlation coefficient is statistically significant is not indicative of a strong association between the two variables. The likelihood of the results of a correlation analysis achieving statistical significance is influenced by study sample size; small samples are unlikely to have sufficient power to detect statistically significant results. Indeed, the results showed non-significant but strong correlations in small studies, and significant but weak correlations in larger studies; the majority of the studies where a strong correlation was detected were small, thus lacking the power to detect significant results. Although pooling these studies to derive a summary estimate could be a solution to increase the statistical power, the high level of clinical heterogeneity across the studies precluded the use of meta-analytical techniques. In addition, a non-significant result reflects only that there is no linear association (which is what the correlation coefficient evaluates) and not that there is no association at all.

These correlation data may be further limited in their usefulness, as BMD is only one factor that impacts on bone strength and therefore on fracture risk,177 and the association between bone turnover markers and BMD may vary depending on the population and on the menopausal status in women.70 There are three factors that determine bone strength: density, architecture and porosity,177 and therefore increases in BMD explain only a limited proportion of the reduction in fracture risk.91,178 Bone turnover impacts on all the three aspects of bone strength,177 and so unlike BMD, biomarkers used to measure bone turnover reflect changes in overall bone strength, and hence potentially fracture risk.91,177 It has been suggested that it is reductions in activation frequencies and reductions in bone turnover resulting in thickening of critical trabeculae, reductions in perforation of trabecular plates and promotion of bone mineralisation, rather than increases in BMD, that reduce a patient’s risk of fracture with antiresorptive therapy.48,179 In studies of raloxifene, risendronate, alendronate and zoledronic acid, bone turnover markers reportedly explained between 28% and 77% of the fracture risk reduction, compared with changes in BMD explaining up to 28% with the same agents.47 Therefore, even if there is no significant correlation between changes in bone turnover markers and increases in BMD, it cannot be assumed that there is no correlation with fracture risk. BMD is a poor surrogate on which to assess the accuracy of bone turnover markers, and any presumption that bone turnover markers are not effective in identifying patients who are not responsive to treatment and still at risk of fracture based on these data would be unwise.70,177,178
The lack of value of the correlation data was compounded by the extreme heterogeneity across the studies in terms of patient populations, the definitions used to diagnose osteoporosis, treatment regimens, bone turnover marker tests used and their timing, and the DXA sites used and the timing of DXA scans; no two studies reported the results for the same combination of these factors, precluding the use of any meta-analytical models. In addition, where reported, there was heterogeneity across studies in terms of obtaining and handling of samples. For the studies evaluating serum bone turnover markers, patients were not fasted in one study using sBALP, storage temperature ranged from −20 °C to −80 °C, and the assay method used also varied (most noticeably for sBALP, which was measured using chemiluminescence, enzyme immunoassay, ELISA, immunocapture, IRMA, or an ostase assay where the variant was not reported). For studies evaluating urinary bone turnover markers, where reported, collection was the first morning void or second morning void, storage temperature ranged from −20 °C to −70 °C, and the assay method used varied, with chemiluminescence and ELISA being used in studies that reported the assay; two of the studies using uNTX did not report whether or not the results were corrected for creatinine. These limitations in the data prevented us from drawing any strong conclusions regarding the relationship between change in bone turnover markers and BMD, or fracture, from the results of these correlation analyses.

Correlations between changes in bone turnover markers and biopsy results and, more importantly, fracture outcomes may be more useful than correlations between changes in bone turnover markers and BMD. These two outcomes do not solely measure changes in bone density, but incorporate bone porosity and microarchitecture. In addition, an assessment of the association with fracture incidence is a direct assessment of the association with the event the bone turnover markers are attempting to predict at the time of testing. The only study to report correlations to both BMD and fracture incidence showed that post-menopausal women over 55 years of age receiving teriparatide reported statistically significant correlations between changes in sP1NP at 1 month and BMD at 24 months, but not fracture. However, the correlations with BMD were weak, and the correlation coefficients for the association with fracture were not reported; therefore, the strength of the association with fracture is unclear. Whether this would be true of other population–bone turnover marker–treatment combinations, or if sP1NP was measured after more than 1 month of treatment, is uncertain.

The assessment of the relationship between changes in bone turnover markers and biopsy or fracture outcomes would be further improved if confounding factors were adjusted for in a multivariate regression analysis. Even if significant correlations are identified between bone turnover markers and these other measures for fracture risk, it is unclear whether or not the use of correlation statistics fully explains the relationship. If these variables were to be incorporated into a multiple regression analysis where other important predictive variables are also included, it is possible that these other variables, or combination of variables, may be stronger predictors of fracture risk than bone turnover markers. In addition, where there is a non-significant association between change in bone turnover markers and fracture risk, the association may change and become significant when other predictive factors are included in a multivariate regression analysis, owing to a synergistic effect of combined variables. These apparently non-significant associations may therefore be important influences within a multivariate regression analysis. Therefore, evaluations of the association between changes in bone turnover markers and subsequent fracture risk outcomes should incorporate confounding factors. In order to assess the accuracy of bone turnover markers in terms of their ability to identify patients who remain at risk of fracture, research needs to be conducted to investigate their independent predictive value of fracture risk.

Data from studies conducting multiple regression analyses gave some indications that a significant association between the changes in bone turnover markers and the incidence of hip or vertebral fractures in patients being treated with bisphosphonates was observed. However, although we have indicated that data from multiple regression analyses that adjust for important confounding factors are of more value than data from correlation analyses, the limited use of these analyses, and the heterogeneity across the studies that did conduct them, limits the usefulness of these data.
As a consequence of the limitations of the data discussed, there is currently insufficient high-quality, consistent evidence available to draw any firm conclusions on the ability of changes in bone turnover markers to identify patients not responding to treatment, or to predict future fracture risk. There are substantial gaps in the clinical evidence base, particularly in terms of the impact of bone turnover marker monitoring on treatment management decisions, and the independent predictive value of bone turnover markers for future fracture. Further research is required. However, given the large number of possible combinations of patient population–treatment–monitoring regimens, decision-analytic modelling will be an essential component of that research if we aim to inform efficient decision-making. Therefore, suggestions for future research are discussed in Chapter 6, Suggested research priorities so that these can incorporate the needs of the assessments of both clinical effectiveness and cost-effectiveness.

Comparison with previous systematic reviews

One review included RCTs that compared antiresorptive treatment regimens with placebo in post-menopausal women, reported associations between changes in a bone turnover marker (or BMD) and the incidence of fracture. The review was based on 18 trials involving 26,494 women, which amounted to 69,369 woman-years. Poisson regression analyses were used to investigate associations between the difference in percentage change in bone turnover marker between the treatment and placebo groups and the relative risk of fracture, weighted by woman-years. As with our review, the trials were heterogeneous in terms of the bone turnover markers used, treatment regimens, patient population and duration of follow-up over which fracture was assessed. Antiresorptive therapy in general resulted in a decrease in resorption bone turnover marker, and a concomitant decrease in the rate of bone formation turnover markers. Therefore, a decrease in either type of marker is seen as an indication of a reduction in bone turnover. Overall, regression coefficients were 0.0067 [standard error (SE) 0.0034; range of change in bone turnover marker compared with placebo across studies +1% to −70%] for changes in resorption bone turnover marker and 0.0134 (0.0051; range of change in bone turnover marker compared with placebo across studies +7% to −56%) for formation markers. There was a 40% decrease in fracture risk for those treatments that decrease resorption markers by 70% compared with placebo and a 44% decrease in fracture risk for those treatments that decrease formation markers by 50% compared with placebo. Changes in bone turnover marker were significantly correlated with changes in BMD (p ≤ 0.002); regression coefficient (\(R^2\)) = 0.58 for resorption markers and 0.41 for formation markers.

It is unclear how comparable the results of this review are to our review, for several reasons. Firstly, trials that evaluated calcitonin and alendronate combined with oestrogen were included in the analyses; these are treatments not being considered in our review. Calcitonin is now authorised only for short-term use in Paget’s disease and acute bone loss due to sudden immobilisation and hypercalcaemia caused by cancer, and not for treatment of osteoporosis, and oestrogen is a HRT used to treat and prevent a range of post-menopausal symptoms which could lead to an overestimation of the effectiveness of alendronate. Secondly, data for resorption markers (urinary deoxypyridinoline, CTX, NTX and urinary hydroxyproline) were combined, as were data for formation markers (serum osteocalcin and sBALP); the specific bone turnover markers used in each study were not reported in the publication, and therefore it is unclear how many of the studies used the bone turnover markers being evaluated in our review. Thirdly, the analyses were restricted to changes in bone turnover markers at 12 months (three studies used data from 6 months where 12-month data were not available); our review was particularly interested in the changes at 3 (and secondarily at 1 and 6) months, as it is the early detection of treatment non-responders that makes bone turnover marker monitoring a potentially useful strategy. In addition, sensitivity analyses conducted by the review authors showed that the overall statistically significant association between resorption and bone formation turnover markers, and the risk of fracture was lost when the largest trial that administered raloxifene was removed from the analysis (p = 0.09), and for resorption bone turnover markers when any one of three alendronate trials were removed, two of which were the second and third largest trials. Therefore, it is uncertain whether there are specific population–treatment–bone turnover marker combinations that produce significant associations with fracture, or whether this is a sample size issue and that most studies were underpowered to detect the association.
A second review that addressed a broader question, including studies of patients with and without osteoporosis, conducted a separate analysis investigating the association between bone turnover markers and fracture outcomes in osteoporotic patients receiving antiresorptive therapy. This section of the review included seven publications; one of these reported only on the results of serum osteocalcin, which is not being evaluated in our review. All six of the relevant studies included in the review by Vasikaran et al. were included in our review. The same data (results from regression analyses) were extracted for both reviews from one study. However, the data extracted from the other studies varied between the two reviews, making comparisons of the results of the two reviews more difficult. Vasikaran et al. extracted the percentage treatment effect explained (TEE) for two of the studies, an outcome not considered in our review; we extracted regression data from both of these studies, whereas Vasikaran et al. extracted regression data from only one. Predictive data in the form of AUCs were extracted by both reviews from one study; we also extracted regression data. Both reviews extracted regression data from a further study; we extracted data only for sBALP, as this is the only bone turnover marker that had results reported for the subgroup of patients with osteoporosis. The analyses of the other bone turnover markers included patients with osteopenia who were not being considered in our review. For the final study included in the review by Vasikaran et al., we extracted data on the correlations between bone turnover markers (sP1NP and sCTX) and fracture outcomes, whereas Vasikaran et al. extracted the relative risk/odds ratio of fracture for the treatment compared with placebo in the subgroup of patients who had been tested using sP1NP. In addition, our review included results from regression analyses from three studies, and other predictive data from three studies, that were not included in the review by Vasikaran et al. The review by Vasikaran et al. did not consider results from correlation analyses; we included data on correlations between bone turnover markers and fracture from two studies.

Vasikaran et al. concluded that the evidence available relating to the association between bone turnover marker changes and fracture risk reduction is promising, but further studies are needed to address sample handling, the timing of bone turnover marker testing, and the statistical methods used; an assessment of whether or not the final bone turnover marker level is a guide to fracture risk was suggested. Vasikaran et al. also included an assessment of the impact of bone turnover monitoring on adherence to osteoporosis medication. Vasikaran et al. included two studies in this assessment, one of which was included in our review and the other excluded as it was conducted on patients with osteopenia rather than osteoporosis. We included a further four RCTs that evaluated the impact of feedback of bone turnover marker results in patients with osteoporosis. Vasikaran et al. concluded that both of the studies included in their review showed that the feedback of results from a positive result encouraged adherence, and a negative result discouraged it. As already mentioned, this response may be surprising; however, the two studies were conducted in populations that were not representative of a general osteoporotic population or of a population of osteoporotic post-menopausal women (one conducted in older women aged 65–80 years and the other in patients who had not yet developed osteoporosis). Therefore, it is unclear whether or not this response would be representative of the osteoporotic population seen in clinical practice. In addition, the inclusion of additional RCTs in our review casts doubt on the impact of feedback of bone turnover marker test results on compliance and adherence. However, as stated above in Limitations of the available evidence, these studies are unlikely to be reflective of clinical practice.

A third review evaluated the use of bone turnover markers for the monitoring of osteoporosis therapy in post-menopausal women. This review included 48 studies, most of which were excluded from our review because they evaluated serum osteocalcin, tartrate-resistant acid phosphatase 5b (TRAPc5b) or P1CP; recruited patients who did not have osteoporosis or mixed populations where results for those with osteoporosis were not reported separately; made comparisons between baseline bone turnover marker test results and BMD or fracture outcomes only; had fewer than 20 treated patients in their analyses; or received an osteoporotic treatment that was not one of those being considered in this review. Of the 48 studies included in Funck-Brentano et al., nine assessed the correlation between changes in bone turnover markers and either BMD or fracture risk in at least 20 patients with osteoporosis.
receiving one of the treatments being considered in this review, and were therefore included in our review.\textsuperscript{132,139,140,151,157,160,161,163,165} We included a further 17 studies that reported correlations between bone turnover markers and BMD or fracture risk in patients with osteoporosis receiving one of the treatment regimens of interest,\textsuperscript{14,38,40,42–44,58,99,106,131,135,145,146,153,155,159,160} two that reported correlations with results from biopsy,\textsuperscript{150,152} and one that reported correlations with vertebral strength index.\textsuperscript{136} Funk-Brentano et al.\textsuperscript{6} did not include data from regression analyses or other types of predictive data. The review by Funk-Brentano et al. concluded that short-term changes in bone turnover markers were significantly correlated with BMD variation, but there was no evidence that they predict benefit on fracture risk at the individual level. A high proportion of the non-significant correlations were in those studies that also met the inclusion criteria for our review and, therefore, if based solely on those nine studies, the conclusions drawn are unlikely to have been so strong.

Summary

There was no evidence available evaluating the clinical effectiveness of treatment monitoring including a bone turnover marker. The evidence available relating to the predictive accuracy, reliability and reproducibility was heterogeneous and of low quality, precluding the ability to draw any conclusions as to which bone turnover marker should be introduced into routine practice for the monitoring of response to osteoporosis treatments in the absence of evidence from RCTs. Much of the available evidence was in the form of correlations between changes in bone turnover BMD. As stated previously, BMD is a poor surrogate for fracture risk on which to assess the accuracy of bone turnover markers, and any presumption that bone turnover markers are not effective in identifying patients who are not responsive to treatment and still at risk of fracture based on these data would be unwise. Further research is required, particularly in terms of the impact of bone turnover marker monitoring on treatment management decisions, and the independent predictive value of bone turnover markers for future fracture risk. However, suggestions for future research need to be made with modelling in mind; therefore, in order to achieve this, recommendations are discussed in the overall discussion (see Chapter 6, Suggested research priorities). No studies met the inclusion criteria for the systematic review of the cost-effectiveness of bone turnover marker monitoring strategies.
Chapter 4 Economic modelling

Methods for the methodological scoping review

In addition to the systematic review, we undertook a number of additional small focused searches of the economic databases searched as part of the systematic review (HEED, IDEAS, and NHS EED) to identify studies that modelled adherence and/or treatment change in patients in the context of monitoring treatment response. The purpose was to identify modelling methods that may be useful for the development of economic models in the future. No critical appraisal was planned. The objective was to survey the modelling methods and provide a narrative synthesis. Data were extracted on the country, type of model, study objective, adherence definition, approach to modelling adherence and the approach to modelling treatment change. Full search strategies for each database searched are provided in Appendix 1.

Paper titles and abstracts were examined for relevance and all potentially relevant papers were ordered. Papers were then screened for relevance by two reviewers, with disagreements resolved by consensus.

Results of the methodological scoping review

The searches identified 130 records after deduplication; 21 papers were retrieved. Of these, 12 modelled adherence to treatment for osteoporosis and one modelled treatment change and adherence, and were included in the scoping review. The search strategy was not limited by indication, but did not identify any studies that modelled adherence and/or treatment change as a result of monitoring treatment response using a biomarker within a different indication. This section provides a narrative summary of the modelling methods for adherence and treatment management used in the 12 included studies. Summary tables for each study can be found in Appendix 5.

Country

The country setting for four studies was North America:167,182–184 three in the USA167,182,183 and one in Canada.184 The remaining eight papers were conducted for western European countries.185–192 Five studies were conducted by the same authors in Belgium,188–192 (these are included separately because there were slight variations in the methods used); one was undertaken in the UK,187 one in Sweden,185 and one in the Netherlands and the UK.186

Type of model

Nine studies were based on individual patient-level Markov (or state transition) microsimulation models.182,185–192 Five of these were by the same group of Belgian investigators.188–192 Two studies used a combined decision tree and Markov modelling approach,167,184 and one used a Markov cohort model alone.183

Study objective

The objectives of the studies were to estimate the cost-effectiveness of different treatments for osteoporosis;186–188,190 to model adherence or persistence to treatments of osteoporosis;182,183,185,189,191,192 to model the cost-effectiveness of a multifaceted intervention to improve osteoporotic care by encouraging patients to come forward and receive treatment after fracture of the wrist;184 and to evaluate different follow-up regimes for antiresorptive treatments for post-menopausal women with osteoporosis.167 The latter assessed only health benefits, measured in QALYs, and did not include costs.
**Adherence definition**
Nine studies used the term ‘adherence’. ¹⁶⁷, ¹⁸⁵–¹⁹² Ten studies used the term ‘compliance’. ¹⁶⁷, ¹⁸⁵–¹⁹²
Nine studies used the term ‘persistence’. ¹⁸³–¹⁸⁵, ¹⁸⁷–¹⁹²

In Chapter 3 (see Assessment of clinical effectiveness) we adopted standard definitions for adherence, compliance and persistence owing to variation in their use. Persistence is the time until discontinuation of medication, compliance is the proportion of medication taken, and adherence is a combination of persistence and compliance. The remaining sections will use these definitions, rather than those specified by the study authors, to ensure consistency. However, where the terms used differ from these definitions, this will be highlighted.

**Compliance**
Ten studies incorporated compliance into their modelling. ¹⁶⁷, ¹⁸⁵–¹⁹² In seven of the studies, cut-off compliance thresholds were applied to delineate compliant from non-compliant patients. In four studies, patients were considered to be compliant if their medical possession ratio (MPR) was \( \geq 80\% \). ¹⁸⁵, ¹⁹⁰–¹⁹² In one study, compliance was defined in terms of the MPR which ranged from 10% to 100%. ¹⁸⁹ In two studies, a compliance rate was used, but the threshold was not stated. ¹⁸⁶, ¹⁸⁸ Three studies assumed that when a patient was on treatment then they were 100% compliant with their medication use. ¹⁶⁷, ¹⁸², ¹⁸⁷

**Persistence**
Eleven studies incorporated persistence into their model. ¹⁶⁷, ¹⁸²–¹⁸⁵, ¹⁸⁷–¹⁹² A definition of ‘primary non-adherence’ was also mentioned in two studies and was used to describe the situation where patients were prescribed a drug but never had the prescription filled. ¹⁸⁵, ¹⁹²

**Compliance and persistence**
Of the studies that modelled persistence, six were studies which modelled compliance using a range of cut-offs between 0% and 100%. ¹⁸⁵, ¹⁸⁸–¹⁹² Other studies that modelled persistence assumed that all those on treatment were fully compliant.

**How compliance, persistence and adherence were modelled**

**Compliance**
In the four studies which defined compliance as a MPR \( \geq 80\% \), the probabilities of a MPR < 80% were modelled with declining annual percentage rates assigned following the first, second and third years of therapy; ¹⁸⁵, ¹⁹⁰–¹⁹² all conducted sensitivity analyses on the compliance rate. The rate of fracture depended on the MPRs, which were derived from Belgian observational data. In the study where the compliance threshold ranged from 10% to 100% based on individual patient data, the risk of hip fracture for the different thresholds was again estimated from Belgian observational data, and the risk of non-hip fracture was estimated from US observational data. ¹⁸⁹ In one study the compliance rate was estimated at 70.5%. ¹⁸⁸ In one study full compliance was assumed in each arm in the base-case analysis, but non-compliance was assumed to be 30% in sensitivity analysis to evaluate the effect of different compliance rates on the results. ¹⁸⁶ Three further studies assumed that when a patient was on treatment then they were 100% compliant with their medication use. ¹⁶⁷, ¹⁸², ¹⁸⁷ Two of these studies conducted sensitivity analyses on the compliance rate. ¹⁶⁷, ¹⁸⁷ In one of these studies fracture risk for those adherent to treatment was further differentiated depending on whether or not the patient was considered a treatment responder or non-responder. ¹⁶⁷

**Persistence**
In related studies by the same authors, persistence was modelled as the percentage of patients who initiated treatment and subsequently discontinued treatment at different time points ranging from 3 months to 2.5 years. ¹⁸⁸–¹⁹² The time points and the discontinuation rates varied between studies. In one study, persistence rates were estimated for the first 3 years and then were assumed to be stable from 3 years until 5 years. ¹⁸⁵ Another study modelled persistence as 39% at 6 months and assumed a continual
In three studies, patients modelled as discontinuing treatment at 3 months were assigned no treatment benefit but incurred 3 months’ drug/treatment and monitoring costs. In four studies, patients modelled as discontinuing treatment at 6 months were assigned no treatment benefit but incurred 3 months’ drug/treatment and monitoring costs.

Compliance and persistence
In every paper judged as distinguishing the different aspects of adherence, compliance rates were applied to patients continuing treatment. In two of these studies primary non-adherence was also incorporated and was modelled as 4.6–11.6%. It was assumed that these patients incurred only the cost of a physician visit and BMD measurement in one study, or cost of screening in another.

How treatment management was modelled
Only one study incorporated a change of treatment into their model structure. The model allowed for switching from a bisphosphonate to a second-line treatment, such as teriparatide, if results of a bone turnover marker test during follow-up led to the conclusion that compliance or response to treatment was inadequate. The authors appeared to have assumed that bone turnover markers were able to correctly identify responders and non-responders to treatment, and so test accuracy data were not included in their model.

Summary of approaches for modelling adherence and treatment change
Only one study modelled treatment change as a result of bone turnover marker monitoring during treatment, and this study did not include test accuracy data. The other studies only modelled aspects of adherence to treatment; compliance was the most commonly modelled variable, with persistence and compliance being distinguished in some studies.

Compliance was considered a binary variable in most studies that modelled compliance. That would allow a relative risk of compliance to be utilised in a model of monitoring strategies that gave test feedback to patients to encourage adherence. The cut-off point for determining compliance/non-compliance was 80% in several studies, chosen to be consistent with that reported in the literature, although it is not clear how that cut-off point was derived. It is likely that there will be more evidence on the risk of fracture for compliant and non-compliant patients defined with that cut-off point than for other cut-off points. Potential confounders should be accounted for when estimating the relative risk of fracture between compliant and non-compliant patients.

Real-life persistence rates were estimated, where possible, based on observational data from databases, and hazard rates could be estimated if survival models fit the data. Six studies modelled compliance, non-compliance and persistence separately, and therefore incorporated the different aspects of adherence.

Including an estimate of primary non-adherence where patients are prescribed treatment is a useful approach if there are a significant proportion of patients for whom that applies. These primary non-adherent patients would not be captured in a survival model.
Economic model

As previously stated, the systematic review of cost-effectiveness analyses did not identify any relevant published studies. In addition, the review of the clinical effectiveness evidence could not establish the clinical effectiveness of bone turnover marker monitoring strategies. As well as a lack of clinical effectiveness data, other key parameters for which there were inadequate data included the test accuracy data at different time points; fracture risk given compliance and responder status; and the effect of bone turnover marker feedback on compliance and persistence for different tests and treatment regimens. Owing to the lack of these relevant data, no economic model was developed and consequently no expected value of information analyses conducted. Expected value of information analyses can be conducted only when valid estimates for the model parameters are available and valid estimates of uncertainty are available for the relevant parameters.

This section, therefore, describes what the necessary information to undertake a cost-effectiveness analysis of the treatment and monitoring strategies described in the section below, might be, and the research that is required to fill the gaps in the current evidence base.

The objective of an economic model is to determine the cost-effectiveness of bone turnover markers for monitoring treatment response and informing patient management decisions. There are two potential uses of bone turnover marker tests within the proposed monitoring strategy: (1) to encourage patients to adhere to treatment; and (2) to inform treatment change. The premise for both uses is that the bone turnover marker tests can accurately identify response to treatment. The decision alternatives relevant to the decision question posed will include various treatment and test combinations, including the option of no treatment. The optimal test will depend on the choice of treatment, which includes the dose, frequency and mode of delivery. The timing of the tests and associated decision rules may depend on the type of test (e.g. P1NP, BALP, CTX) and the threshold cut-off point for determining treatment response and non-response. Decision rules involve specifying a timetable of tests and the patient pathway that would be followed based on the results of those tests. Having defined these alternatives, the cost-effectiveness of the alternative monitoring strategies can be assessed.

Current practice

Based on discussion with a clinical advisor the timings of GP visits and DXA tests in current practice are presented in Figure 3. In this monitoring strategy decisions on treatment change as a result of a poor DXA test result occur at either 2 or 3 years, depending on whether or not a fracture occurred within the first 2 years. Decisions may also be influenced by the occurrence of side effects and patient statement on compliance to treatment: if a patient discontinues treatment because of side effects, then a different class of treatment may be prescribed. If there is a poor DXA test result and the patient claims to be compliant to treatment then a new class of treatment would be prescribed; however, if the patient admits to non-compliance then an alternative dose or mode of delivery of the same treatment may be appropriate, where one is available.

Alternative monitoring strategy using bone turnover marker tests

We are not aware of any published guidelines on the use of bone turnover markers in the UK for monitoring response to treatment for osteoporosis which could be used to inform alternative monitoring strategies incorporating bone turnover markers. Therefore, based on discussion with a clinical advisor, a feasible monitoring strategy using bone turnover marker tests was constructed and is presented in Figure 3. The purpose of this figure is to allow the presentation and discussion around the type of data that might be required to evaluate similar strategies.

The numbers in circles indicate four different stages to the strategy. In this hypothetical strategy, patients are started on treatment at stage 1. At stage 2, after 3–6 months, depending on the type of test and treatment, a first bone turnover marker test is done. The results are fed back to the patients in order to
A: Current monitoring strategy

- Start Tx
- BM test and GP visit
  - 3 months
- DXA test and GP visit
  - +ve result: Same Tx
  - -ve result: Change Tx
  - 2 years if fracture occurred/3 years if no fracture occurred by 2 years

B: Feasible monitoring strategy

- Start Tx
- BM test and GP visit
  - 3 months
- +ve result: Same Tx
- -ve result: Change Tx

1. Same Tx
- BM test and GP visit
  - -ve result: Change Tx
  - 6 months
- +ve result: Same Tx
- DXA test and GP visit
  - -ve result: Change Tx

2. Change Tx
- DXA test and GP visit
  - -ve result: Change Tx
  - +ve result: Same Tx
- 2 years if fracture occurred/3 years if no fracture occurred by 2 years

3. Change Tx
- DXA test and GP visit
  - -ve result: Change Tx
  - +ve result: Same Tx
- 2 years if fracture occurred/3 years if no fracture occurred by 2 years

4. Change Tx

FIGURE 3 Current and feasible monitoring strategies for response to treatment. a, GP visit: enquire about patient adverse events and adherence. The numbers in circles mark different stages of the monitoring strategy. Tx, treatment; +ve, positive; -ve, negative.
improve, or encourage continued, adherence. At stage 3, a second bone turnover marker test is done only for patients who received a negative result on the first test. This would be done at 6–12 months depending on the type of test and treatment. Using a predefined decision rule after a second negative test result, a treatment change would be recommended. The clinician may recommend a new treatment class to everyone with two negative test results. Alternatively, if the clinician has assessed patient adherence, he or she may recommend a new treatment class only to adherent patients and a different mode of delivery and/or frequency of the same treatment class to non-adherent patients where one is available.

Given the lack of evidence, it has not been possible to determine clinically relevant strategies for evaluation. The timing of the bone turnover marker tests may vary by the type of test used and the type of treatment administered that the test is monitoring a response to. The optimal timing of a bone turnover marker test needs to be determined by, for example, studies that evaluate the S/N ratio at different time points; there may be a trade-off between a greater S/N ratio and delay to response monitoring. While some S/N ratio data were identified in the review, these were inadequate to address the issue of optimal timing and combination of treatment and tests.

Different cost-effectiveness analysis methods
Based of the review of modelling methods, the two most common models employed to address this type of decision question are:

(a) individual patient-level Markov simulation
(b) decision tree and Markov model combined.

Both of these modelling methods require the combination of evidence on the risk of fracture given different risk factors, the effect of feedback on response rates and bone turnover marker test accuracy data.

Modelling data requirements for hypothetical strategy

Stage 1: osteoporosis treatment
At the start of the model a population will be defined and individuals will be assigned characteristics. For this discussion our population will comprise post-menopausal women. Each woman will belong to an age group, will have an assigned T-score (e.g. < –2.5) and a fracture history (e.g. no prior fractures or two prior fractures). She will start on treatment for osteoporosis. Based on her characteristics (e.g. T-score < –2.5, no prior fractures, aged 55 years) and treatment (e.g. oral bisphosphonates) there will be an associated fracture risk. That risk will vary depending on whether a patient is a responder or non-responder to treatment.

The ideal method to model the true-positive, false-positive, true-negative and false-negative test results, and their impact on treatment change later on in the model, is to distinguish patients by treatment response defined by a cut-off point related to the test being used. The fracture risk for these population groups would need to be determined. The optimal cut-off point in terms of the least significant change for any of the bone turnover markers is uncertain.

Stage 2: initial monitoring stage
The second stage of the model occurs when the first bone turnover marker test is done. In our hypothetical strategy this is at 3 months, but in clinical practice this would be a variable optimal time point determined for each test and treatment combination being evaluated. Bone turnover markers test for treatment response. In the hypothetical strategy we have assumed that poor results of the first bone turnover marker test feedback will be a signal to the clinician to encourage compliance and persistence.

Stage 3: the effect of bone turnover marker tests on treatment change
Stage 3 of the model is when a second bone turnover marker test is done to identify response to treatment. At this stage clear decision rules regarding what defines a responder from a non-responder in
terms of bone turnover marker test results would be required. Only patients with a negative result from the first test get a second test. The decision rule considered in the hypothetical strategy is that a change in treatment class would occur only if there were two negative test results.

Stage 4: the effect of a dual-energy X-ray absorptiometry test on the treatment pathway

At stage 4 of the hypothetical strategy DXA testing is conducted at 2 or 3 years after the start of treatment. DXA testing is also done in current practice at 2 or 3 years.

Test accuracy

When diagnostic or prognostic tests are included in a decision-analytic model, test accuracy needs to be considered and the clinical outcomes for patients with correct (true-positive and true-negative) and incorrect (false-positive and false-negative) test results need to be incorporated.\(^\text{193}\) Test accuracy will be an important factor in any decision model used to investigate the current decision problem as test errors may result in an incorrect treatment pathway being followed. A non-responder incorrectly identified as a responder (false-positive) would likely remain on their current treatment, and would not benefit from a change in treatment class. Conversely, a responder who was incorrectly classified as a non-responder (false-negative) may be prescribed a different drug unnecessarily. Where patients have true-positive (a responder identified as a responder) or true-negative (a non-responder identified as a non-responder) results, they will benefit from appropriate treatment management choices based on the correct test results.

For this decision problem, test accuracy will need to be determined for each of the tests within the pathway; in the example, test accuracy for bone turnover markers would be needed for stages 2 and 3, and test accuracy for DXA at stage 4. The major limitation for establishing the accuracy of bone turnover marker tests in identifying treatment non-responders is the lack of a gold standard.

There are a range of reference standards available, but each has its limitations. Bone biopsy could be seen as the ideal reference standard. However, given the invasiveness of the test and the high risk of complications it would be considered unethical to conduct studies where all patients would undergo biopsy to confirm treatment response. The next most complete reference standard would be the use of BMD with clinical follow-up to determine fracture incidence to supplement the BMD results. Where BMD indicated a response to treatment sufficient enough to reduce a patient’s fracture risk but the patient went on to have a fragility fracture, the patient could be reclassified as a non-responder. This would mean that if the bone turnover marker was negative in this patient, it would be reclassified from a false-negative to a true-negative. Whether a patient who did not have a response in BMD but did not go on to have a fragility fracture within a certain time frame could be reclassified from a non-responder to a responder is less certain.

Less perfect reference standards are the incidence of fracture alone and BMD alone, the latter being the poorest available reference standard. The lack of suitability of BMD as a surrogate for fracture risk against which we can judge the accuracy of bone turnover tests is discussed in Chapter 3 (see Limitations of the available evidence). An alternative method for estimating the test error rates for a test at a particular time point would be to use the intraindividual variation data that is used to calculate S/N ratios, but these estimates may be inferior to test accuracy data with an appropriate comparator.\(^\text{194}\) When establishing the accuracy of DXA, the only reference standards are bone biopsy and fracture incidence, and the limitations of these have already been discussed.

Once the proportion of true-positives, true-negatives, false-positives and false-negatives had been established, fracture risk for each of these groups would need to be estimated.

Measures of adherence

As discussed previously, adherence has two aspects: compliance and persistence – both of these are important causes of non-response to osteoporosis treatment. Clearly, adherence to treatment will affect a
patient’s response to treatment, at all stages of the treatment pathway. Therefore, any attempt to model adherence would apply to each stage in Figure 3. Several studies modelling the effectiveness of treatment have modelled compliance and persistence. Our decision problem involves the ability of tests to identify treatment responders and non-responders rather than to establish treatment effectiveness. For patients identified by a test as a non-responder (bone turnover markers at stage 2 and 3; DXA at stage 4), treatment non-response could have a number of causes, including: non-compliance; non-persistence; an underlying, untreated cause of the osteoporosis; an inability to absorb the drug; and/or test error. The issue of test accuracy has been addressed in earlier in this chapter (see Test accuracy).

Including measures of adherence in the model would be of use only if there would be different treatment pathways for adherent and non-adherent non-responders, and/or there was evidence that feedback of bone turnover marker test results increased adherence. It is possible that there would be different treatment pathways for adherent and non-adherent non-responders for those who were treated with first-line oral bisphosphonates, as true non-responders are unlikely to respond to a different dosing frequency or mode of administration, but patients who are non-compliant with a daily or weekly oral dose may become compliant with a monthly oral dose or intravenous administration. Therefore, the importance of modelling adherence may differ depending on the choice of first-line treatment. Adherence rates may improve by providing feedback from bone turnover marker tests to the patients, but there is no strong evidence for this effect.

The ideal method to model adherence would be to include adherence from the start of treatment. A primary non-adherence estimate could be modelled, which is a percentage of patients who are recommended treatment but do not start treatment. This could be estimated from the number of prescriptions that are not filled; however, this would underestimate the value, as some patients may have a prescription filled and not commence treatment. To include measures of adherence in subsequent stages of a model, the proportions of non-compliant non-responders and compliant non-responders would need to be established. There are limitations with the methods available for determining these estimates. It is likely that the most common means of identifying these data would be based on patient reporting, which is prone to bias. Measures such as the MPR rate are also problematic, as patients may have their prescriptions filled but not take all of the medication prescribed. In the case of bisphosphonates, patients may take their tablets but not comply with the strict regimen required, adding a further complication to the establishment of a rate of compliance. Persistence could be estimated by reviewing the attrition rates from RCTs evaluating treatment effectiveness; this would likely underestimate the rate. Such a review could be supplemented with data from observational studies that use sources such as GP databases. Any estimate of compliance, persistence or overall adherence would be highly uncertain, making it extremely important that it is incorporated into the modelling in a manner which allows appropriate sensitivity analysis to be undertaken.

In addition to the adherence data, the fracture risk for these population groups would also need to be determined. If it can be assumed that the relative risk of fracture given treatment versus no treatment is based on a compliant population, then the relative risk of fracture given good compliance versus poor compliance enables the calculation of the risk of fracture for poorly compliant patients to treatment. There would be an interaction between compliance and response. The optimal cut-off point is uncertain and needs to be established prior to modelling.

One aspect of adherence deserves separate consideration, namely discontinuation rates due to treatment-related adverse events; this is a particular issue for bisphosphonates. Gastrointestinal upsets are common with oral bisphosphonate therapy, but an additional, serious, side effect would have to be incorporated in the model; osteonecrosis of the jaw (ONJ). This is a side effect specific to the use of bisphosphonates, and can lead to the need for costly dental surgery. Although this is a rare adverse event, there is some evidence that the risk of developing ONJ is much higher with intravenous bisphosphonates than with oral preparations. The rate of ONJ may be investigated initially by a review of long-term RCTs.
that evaluate the clinical effectiveness and safety of bisphosphonate therapy; the review could supplement these data with the available prospective observational and retrospective studies.

**Summary of modelling approaches and available evidence**

The key part of a cost-effectiveness analysis of bone turnover marker tests for monitoring response to treatment for osteoporosis is accounting for test accuracy, the prognostic outcomes for true- and false-positive and negative test outcomes, and the effect of bone turnover marker feedback on patient adherence to treatment. This affects both who benefits from bone turnover marker feedback and who benefits from treatment change. These data were either absent completely in the evidence identified in this review, were insufficient given the different tests and treatments, or were applicable to populations with unrealistic adherence rates for clinical practice.

Compliance and persistence are commonly modelled separately. The effect of test feedback on adherence is often reported as the effect on compliance or persistence. Given the relationship between the two and bone turnover marker levels, this does require assumptions to simplify the model.

This section has focused on those gaps in the evidence that would be essential to any future decision-analytic model but may be difficult to establish estimates for. Other variables, such as estimates of treatment effectiveness and data on utilities or other relevant QoL outcomes, would be required. Information would also be required on resource use, including the cost of tests, GP/clinic visits, treatment costs, and the costs associated with treating serious side effects such as ONJ.
Chapter 5 Discussion

Statement of principal findings

The systematic review of clinical effectiveness found no evidence evaluating the impact of treatment monitoring regimens that included a relevant bone turnover marker on treatment management decisions. The review identified limited data assessing the effect of bone turnover marker feedback on patient compliance, persistence and/or adherence to treatment, the results of which suggested that the positive feedback results encouraged patient adherence.

A moderate number of correlation data were identified relating to the predictive accuracy of the four bone turnover markers, namely P1NP, BALP, CTX and NTX, in osteoporotic patients being treated with one of the targeted drug therapies: bisphosphonates, raloxifene, strontium ranelate, teriparatide and denosumab. Most correlations had a small effect size, indicating a weak association between changes in the level of bone turnover markers (usually between 1 month and 6 months of starting treatment) and subsequent changes in BMD (usually between 1 year and 3 years after the start of treatment). The studies that used regression analyses to adjust for confounding factors when evaluating the association between bone turnover markers and either subsequent changes in BMD or fracture gave some indication that changes in bone turnover markers may be significantly associated with these outcomes; however, there were too few of these studies to draw any firm conclusions.

In terms of the evaluation of the test reliability and reproducibility, some available evidence suggested that sP1NP may have a greater S/N ratio than sBALP and sCTX at short-term follow-up, but the data on this outcome were sparse and longer-term follow-up data absent.

Overall, the evidence required to address the decision problem was lacking, and the limited evidence that was available was heterogeneous and of poor quality. Consequently, it was impossible to draw any conclusion as to whether bone turnover markers were able to identify treatment non-responders or predict fracture risk independently of BMD in patients receiving osteoporosis treatment.

The systematic review of cost-effectiveness identified no studies evaluating different treatment monitoring strategies, where BALP, P1NP, CTX or NTX were incorporated as part of one of the strategies. Given the lack of evidence on the clinical effectiveness of bone turnover marker monitoring on treatment management, a de novo decision-analytic model could not be developed, and consequently the value of any future research could not be investigated.

The scoping review of modelling methods used in the broader context of osteoporosis treatment identified 12 modelling studies, of which only one modelled treatment change. A range of methods were used to deal with modelling adherence; adherence was defined in different ways, but several studies distinguished between compliance and persistence components of adherence which was consistent with the standard definitions that this review adopted from Delmas.56 Compliance was defined using the MPR, and the threshold for distinguishing compliant and non-compliant patients varied, although 80% was the most common threshold. Real-life persistence rates based on observational data were often modelled for different time points, and primary non-adherence, where patients fail to start treatment, was also included in two studies. Separating these different components of adherence presents a practical method of modelling adherence and the effect of adherence on fracture risk. The one model that incorporated treatment change allowed for switching if the clinician concluded that compliance or response to treatment was inadequate, but the authors assumed perfect test accuracy.
**Strengths and limitations of the assessment**

**Strengths**
We conducted a rigorous systematic review which addressed clear research questions using pre-defined inclusion criteria. Comprehensive literature searches were conducted to locate all relevant published and unpublished studies without any language restrictions, thereby minimising both publication and language biases. Efforts were also made to identify additional studies by hand-searching the reference lists of relevant publications. We are therefore confident that we have been able to include all the relevant studies in the evaluation of the clinical effectiveness, predictive accuracy, reliability and reproducibility, and cost-effectiveness of bone turnover markers for monitoring treatment response that met our inclusion criteria. Each stage of the review was conducted in duplicate, reducing the risk of error and bias. Owing to the high degree of clinical heterogeneity across the studies in terms of patient populations, treatment regimens and duration of follow-up, the clinical data were appropriately synthesised using a narrative approach.

As we were unable to conduct a cost-effectiveness analysis, we conducted additional assessments of the economic models used to address similar decision problems in order to inform future decision-analytic modelling that may be undertaken to address the decision problem investigated in this review.

**Limitations**
The main limitations of the review of clinical effectiveness were the lack of data on the effectiveness of monitoring regimens and the poor quality of the data that were available; all of the included studies were judged as low quality. This lack of robust data relating to the comparative predictive values of the different bone turnover marker tests for monitoring the response to a specific osteoporosis treatment precludes the possibility of determining which bone turnover marker, if any, is superior in terms of its ability to identify treatment non-responders and predict future fracture risk when used for monitoring osteoporosis treatment.

Despite the moderate amount of correlation data identified for the evaluation of the relationship between changes in bone turnover markers and BMD, it remains unclear whether or not these findings can also be utilised as an indicator of the association with the outcome of fracture risk. As discussed in Chapter 3 (see Limitations of the available evidence), BMD is a poor surrogate for fracture risk. Treatment-induced changes in BMD account for a limited proportion of the observed reduction in fracture risk, therefore, using BMD as a surrogate for assessing the accuracy of bone turnover markers for identifying patients on treatment who remain at risk of fracture is inappropriate.

Although the incidence of fracture is a more robust outcome measure for fracture risk, there was a paucity of data correlating the changes in bone turnover marker levels to this outcome. There was also a paucity of studies that adjusted the associations between changes in bone turnover markers and other changes in BMD, fracture, or other measures such as spinal strength indices or results of biopsy, for confounders in multiple regression analyses. As discussed in Chapter 3, where strong associations are identified between bone turnover markers and fracture or other measures in correlation analyses, there is no evidence that these would produce significant associations when other important confounding variables are included in regression analyses. Alternatively, where there is a non-significant correlation between change in bone turnover markers and one of these outcomes, the association may change and become significant when other predictive factors are included in a multivariate regression analysis. Either way, assessing the association between changes in bone turnover markers and any outcome in isolation, without adjusting for confounders within a multivariate regression analysis, is unlikely to reflect the true association between these variables within a patient; it was impossible to draw any conclusion as to whether or not these bone turnover markers were able to identify treatment non-responders and predict fracture risk independently of BMD measurements in patients receiving osteoporosis treatment.

It should be noted that the results from the studies utilising correlation and regression analyses were inconsistent. This may be due to the considerable clinical heterogeneity across the included studies in terms of the definitions used to identify those with osteoporosis for inclusion in the studies, patient
populations recruited, the treatment regimens administered, and the type and timing of the tests being evaluated. Most of the included studies had small sample sizes, resulting in low statistical power to detect a significant association.

The analysis of test reliability and reproducibility in women being treated for osteoporosis was limited; very few studies reported these data. Test reliability and reproducibility is most commonly measured in either healthy individuals or control subjects who are not receiving treatment. Although this provides baseline intraindividual and interindividual CVs and a S/N ratio, it does not inform us as to how these tests perform in women receiving treatment. Whether receiving treatment would increase or decrease these measures of variability is unknown.

The lack of any published decision-analytic models investigating the decision problem being addressed in this review, and the lack of evidence on the effectiveness of monitoring treatment response using bone turnover markers, meant that the cost-effectiveness of different monitoring strategies could not be investigated at this time. In order to construct such a model, the large gaps in the current evidence base will need to be filled. We identified these gaps and the data required for a future cost effectiveness analysis of different monitoring strategies and discussed how these data could be obtained in Chapter 4 (see Economic model). The uncertainties that remain and research priorities are highlighted in the section below and in Chapter 6 (see Suggested research priorities), respectively.

**Uncertainties**

There are currently large gaps in the evidence base relating to the use of bone turnover markers for monitoring osteoporosis treatment. These include:

- the ability of changes in bone turnover markers to identify treatment non-responders
- the ability of changes in bone turnover markers to impact on compliance, persistence and adherence to each of the treatments being evaluated
- the accuracy of changes in bone turnover markers to predict future fracture risk
- the ability of bone turnover markers to inform treatment change
- the most appropriate timing of the conduct of bone turnover marker testing; this may vary depending on the treatment–test combination
- which bone turnover marker is superior in terms of its ability to identify treatment non-responder and predict fracture risks for monitoring specific osteoporosis treatments
- the reliability and reproducibility of bone turnover marker tests in patients receiving treatment for osteoporosis
- the most cost-effective monitoring regimen for patients being treated with bisphosphonates, raloxifene, strontium ranelate, teriparatide or denosumab.
Chapter 6 Conclusions

Implications for service provision

The lack of evidence of clinical effectiveness, and the heterogeneity and poor quality of the available evidence on the accuracy, reliability and reproducibility of bone turnover markers for monitoring response to osteoporosis treatment, precluded the possibility of making any recommendations on the choice of bone turnover marker being used in routine clinical practice for its superiority to monitor osteoporosis treatment response. In addition, the evidence to support the use of bone turnover marker feedback results to improve patient adherence to osteoporosis treatment was not convincing.

Suggested research priorities

In order to determine whether or not bone turnover marker monitoring improves treatment management decisions and ultimately impacts on patient outcomes in terms of reduced incidence of fracture, RCTs are required. The predictive accuracy of bone turnover markers for future fracture outcomes in patients receiving osteoporosis treatment could be investigated using prospective, long-term observational studies with large sample sizes. However, as the nature of bone turnover marker response is determined by the mechanism of action of the drug, any future research needs to identify the most appropriate treatment–test combinations in order to identify whether or not the predictive accuracy of a particular bone turnover marker can be maximised to aid treatment management decisions. All future studies should adopt a standardised definition of osteoporosis, such as the WHO criteria.7

There are potentially a large number of patient population–treatment–test combinations; therefore, conducting RCTs or even larger observational studies to establish the effectiveness for all of these combinations would not be feasible. Therefore, it is likely that identifying the most promising combinations would be beneficial in order to ensure that the most promising are evaluated in the more costly and time-consuming studies such as RCTs. This would include not only identifying which bone turnover test best identifies non-responders to specific treatments, but also the optimal timing of these tests. This may feasibly be achieved through the use of a patient registry, where relevant pre-specified standardised data would be collected and trends both in the use of different tests and in outcomes related to test–treatment combinations could be identified. However, without more widespread use of these tests in clinical practice, the usefulness of such a registry would be questionable. If a registry is not established, a survey of the current use of bone turnover markers may be useful. An alternative to establishing a registry to identify promising test–treatment combinations might be to undertake smaller, less costly studies that identify those treatment–bone turnover marker combinations that have the lowest inter- and intraindividual patient variability (and therefore a higher S/N ratio). These smaller feasibility studies could be used to help to identify the most promising combinations for future more costly research. Such studies would be of use only when conducted in the context of establishing inter- and intraindividual patient variability in bone turnover markers and would not be useful if the decision was made to either establish a patient registry, or if there was considered to be sufficient experience within the clinical setting already available to identify treatment-test combinations that could be evaluated in effectiveness studies.

To further limit the number of RCTs and other costly and time-consuming research, there are some areas of uncertainty that could be classified as low priority. These could be investigated initially within a decision-analytic framework, once sufficient evidence becomes available on monitoring effectiveness and the predictive value of bone turnover markers. By using this strategy, those areas of uncertainty that are key drivers of cost-effectiveness can be identified. Further research can then focus on investigating the impact of those areas of uncertainty that most influence the cost-effectiveness of the monitoring regimens, rather than being conducted to inform those estimates and assumptions to which the cost-effectiveness analysis is robust.
We consider that the research priority is to identify the most promising treatment–test combinations. This can be achieved either by conducting small variability studies or, if more widespread use is feasible, by initiating a patient registry to collect data. The former would be quicker, easier and less costly, but the quality of the data would be poorer and likely to be collected in small selected populations. This would mean that the results may not reflect the broader population seen in clinical practice, and the choices made as to which treatment–test combinations to evaluate in a RCT may be inappropriate. Once the most promising treatment–test combinations have been identified, well-designed RCTs can be conducted to evaluate the clinical effectiveness of those monitoring regimens; this would include measuring outcomes such as the proportion of non-responders, adherence rates, treatment management decisions, and fracture outcome. Data from these RCTs along with other sources can then be included in a decision-analytic model in order to investigate cost-effectiveness.
Acknowledgements

We would like to thank Professor John Kanis for providing clinical advice during the early stages of the protocol development.

Contributions of authors

Dr Jane Burch: researcher responsible for study selection, data extraction, validity assessment, narrative synthesis of the clinical evidence and writing the clinical sections of the report, and providing comments on drafts of the economic sections of the report.

Mr Stephen Rice: researcher contributing to study selection, data extraction, validity assessment, data synthesis, responsible for writing the economic sections of the report and providing comments on drafts of the clinical sections of the report.

Dr Huiqin Yang: researcher contributing to study selection, data extraction, validity assessment, data synthesis and writing clinical sections of the report, and providing comments on drafts of the economic sections of the report.

Ms Aileen Neilson: researcher contributing to study selection, data extraction, validity assessment, data synthesis and writing economic sections of the report, and providing comments on drafts of the clinical sections of the report.

Mrs Lisa Stirk: developed search strategies and conducted a range of searches to locate studies, wrote the methodological sections relating to the search and provided comments on drafts of the report.

Professor Roger Frances: provided clinical advice throughout the project and comments on drafts of the protocol and report.

Dr Paul Holloway: provided clinical advice and comments on drafts of the protocol and report.

Professor Peter Selby: provided clinical advice and comments on drafts of the report.

Ms Dawn Craig: took overall managerial responsibility for the project, contributed to all aspects of the project, and provided comments on drafts of the report.
References


30. antibodies-online.com. *Carboxyterminal Propeptide of Type 1 Procollagen (P1CP)*. URL: www.antibodies-online.com/kit/366576/Carboxyterminal+Propeptide+of+Type+1+Procollagen+P1CP+EUSA (cited August 2012).


63. PRISMA: transparent reporting of systematic reviews and meta-analyses. URL: www.prisma-statement.org (cited 16 March 2011).


66. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions J Epidemiol Community Health 1998;52:377–84. http://dx.doi.org/10.1136/jech.52.6.377


72. Stepan JJ. Clinical utility of bone markers in the evaluation and follow-up of osteoporotic patients: why are the markers poorly accepted by clinicians? J Endocrinol Invest 2003;26:458–63.

REFERENCES


REFERENCES


144. Roche. A randomized open-label study to investigate the impact of Bone Marker Feedback on adherence to monthly oral Bonviva in women with post-menopausal osteoporosis. URL: www.roche-trials.com/trialDetailsGet.action?studyNumber=ML19930 (accessed September 2012).


159. Lane NE, Sanchez S, Genant HK, Jenkins DK, Arnaud CD. Short-term increases in bone turnover markers predict parathyroid hormone-induced spinal bone mineral density gains in postmenopausal women with glucocorticoid-induced osteoporosis. Osteoporos Int 2000;11:434–42. http://dx.doi.org/10.1007/s001980070111


REFERENCES


394. Reid DM. Can high bone turnover markers identify osteopenic postmenopausal women at risk of future fracture? *Nat Clin Pract Endocrinol Metab* 2007;3:570–1. [http://dx.doi.org/10.1038/ncpendmet0560](http://dx.doi.org/10.1038/ncpendmet0560)


396. Lane N, Yao W, Arnaud CD. The association of serum RANKL and OPG levels with other biochemical markers of bone turnover in glucocorticoid induced osteoporosis patients treated with hPTH (1-34). *J Bone Miner Res* 2002;17:S209.


REFERENCES


REFERENCES


Appendix 1 Literature search strategies

MEDLINE and MEDLINE(R) In-Process & Other Non-Indexed Citations (Ovid)

Date range: 1946–February week 4 2012.

Date searched: 7 March 2012.

Records found: 1733.

Search terms

1. (P1NP or PINP).ti,ab. (593)
2. (procollagen adj3 propeptide).ti,ab. (429)
3. (procollagen adj3 peptide).ti,ab. (557)
4. (collagen adj3 propeptide).ti,ab. (191)
5. (BSAP or BALP or BAP).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (1090)
6. bone specific alkaline phosphatase$.ti,ab. (1196)
7. bone alkaline phosphatase$.ti,ab. (1135)
8. bone source alkaline phosphatase$.ti,ab. (1)
9. (CTX or NTX).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (2032)
10. crosslaps.ti,ab. (296)
11. (telopeptide$ adj3 collagen).ti,ab. (668)
12. (n-telopeptide$ adj3 collagen).ti,ab. (214)
13. (c-telopeptide$ adj3 collagen).ti,ab. (184)
14. bone turnover marker$.ti,ab. (1333)
15. bone metabolic marker$.ti,ab. (190)
16. Biological Markers/ and exp "Bone and Bones"/ (3520)
17. ((biochemical marker$ or biomarker$ or biological marker$) adj2 bone$).ti,ab. (2261)
18. bone marker$.ti,ab. (1437)
19. or/1-18 (10656)

Line 19 captures bone turnover marker terms

20. exp osteoporosis/ (39,832)
21. osteoporos$.ti,ab. (43,304)
22. 20 or 21 (55,772)

Line 22 captures osteoporosis terms

23. diphosphonates/ or alendronate/ or clodronic acid/ or etidronic acid/ (15,156)
24. (bisphosphonate$ or diphosphonate$).af. (17,117)
25. (alendronate or alendronic acid or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or mxtral or oseotenk or osteofene or osteonate or phostarac or regenesis or silidral or tilios or bonalen or cleveron or endronax or minusorb or novic or osdron or ossomax or ostenan or osteofar or osteoform or osteoral or osteotrat or recalfe or terost or aldrox or fosval or holadren or leodrin or oseotal or osteofem or osteosan or alefos or fosalan or ostalert or bifosa or osteofos or fosalan or maxbone or adronat or alendros or dronal or genalen or onclast or drovitan or landrofen or sinfackt or aldonrac or ailot or denfos or fixopan or genalmen or osteodur or porosal).af. (4270)
26. (clodronate or clodronic acid or bonefos or loron or clodron or lodronat or ascredar or ostac or clastoban or lytos or clasteon or climaclod or clodeosten or clody ordifosfonal or dolkin or moticlod or nikklo or ossiten or osteonorm or osteostab or mebonat).af. (1827)
27. (etidronate or etidronic acid or EHDP or didronel or difosfen or detidron or osteodidronel or bonemass or ostepro or didronate or diphos or anfozan or biotredine or dralek or etidron or etiplus or feminoflex or maxibral or oflocin or osfo or ostedon or osteodrug or osteoton or ostogene or somaflex or sterdome or svirox or tilferan or monate-os or etidrate or osteum).af. (7341)
28. (ibandronic acid or ibandronate or boniva or bondronat or bonviva or bandrobon or bonat).af. (766)
29. (pamidronate or pamidronic acid or aredia or ADP sodium or aminomux or pandosa or pandrat or pampil or pamit or osteopam or aredrone or pamidria or pamidrane or linoten or xinsidona).af. (2470)
30. (risedorane or risedronic acid or actionel or ductonar or ribastamin or ridron or risedon or optinate or risosof or benet or acrel).af. (2039)
31. (zolecodronic acid or zoleidonate or zometa or zomera or aclasta or reclast or zoldria).af. (2319)
32. (tiludronic acid or tiludronate or skelid).af. (139)
33. (neridronic acid or neridronate or nerixia).af. (63)
34. (olpadronic acid or olpadronate).af. (73)
35. (cimadronic acid or cimadrone).af. (92)
36. (piridronic acid or piridronate).af. (0)
37. (icandronic acid or icandrone or bisphonal).af. (1)
38. (minodronic acid or minodronate).af. (57)
39. (raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax or estroact or ralista or celvista).af. (2870)
40. Raloxifene/ (2063)
41. (strontium ranelate or proteles).af. (421)
42. Teriparatide/ (1167)
43. (denosumab or prolio or xgeva).af. (504)
44. (teriparatide or forteo or forsteo).ti,ab. (587)
45. (treatment$ or treat or treated or treats).ti,ab. (3,156,022)
46. dt.fs. (1,504,349)
47. or/23-46 (3,906,922)

**Line 47 captures intervention terms**

48. 19 and 22 and 47 (1879)

**Line 48 combines bone turnover marker, osteoporosis and intervention terms**

49. exp animals/ not humans/ (3,667,503)
50. 48 not 49 (1733)

**Line 50 excludes animal-only studies**

Key
- / = indexing term [medical subject heading (MeSH) heading]
- exp = exploded MeSH heading
- $ = truncation
- .ti,ab. = terms in either title or abstract fields
- .af. = terms in all fields
- .fs. = floating subheading – searches all MeSH terms which use that subheading
- adj = terms adjacent to each other (same order)
- adj2 = terms within two words of each other (any order)
- adj3 = terms within three words of each other (any order)
Cumulative Index to Nursing and Allied Health Literature (EBSCOhost)

Date range: 1982–date.

Date searched: 7 March 2012.

Records found: 155.

Search terms
S48 S21 and S24 and S47 (155)
S47 S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32 or S33 or S34 or S35 or S36 or S37 or S38 or S39 or S40 or S41 or S42 or S43 or S44 or S45 or S46 (330,229)
S46 treatment* or treat or treated or treats (327,921)
S45 teriparatide or forteo or forsteo (102)
S44 denosumab or prolia or xgeva (108)
S43 "strontium ranelate" or protelos (41)
S42 (MH "Raloxifene") (441)
S41 raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax or estroact or ralista or celvista (636)
S40 "minodronic acid" or minodronate (1)
S39 "icandronic acid" or icandronate or bisphonal (0)
S38 "piridronic acid" or piridronate (0)
S37 "cimadronic acid" or cimadronate (0)
S36 "olpadronic acid" or olpadronate (3)
S35 "neridronic acid" or neridronate or nerixia (4)
S34 "tiludronic acid" or tiludronate or skelid (4)
S33 "zolendronic acid" or zoledronate or zometa or zomera or aclasta or reclast or zoldria (367)
S32 risdonenate or "risedronic acid" or actonel or ductonar or ribastamin or ridron or risedon or optinate or risofos or benet or acrel (238)
S31 pamidronate or "pamidronic acid" or aredia or ADP sodium* or aminomux or pamdosa or pandrat or pamisol or pamitor or osteopam or aredronet or pamidria or pamifos or pamidran or linoten or xinsidona (223)
S30 "ibandronic acid" or ibandronate or boniva or bondronat or bonviva or bandrobon or bonat (109)
S29 etidronate or "etidronic acid" or EHDP or didronel or difosfen or detidron or osteoidronel or bonemass or osteotop or didronate or diphos or anfozan or biotredine or dralen or etidron or etiplus or feminoflex or maxibrin or oflicin or osfo or ostedron or osteodrug or osteoton or ostogene or somaflex or sterodome or sviroxit or tilferan or dionate-os or etidrate or osteum (172)
S28 clodronate or "clodronic acid" or bonefos or loron or clodron or lodronat or ascredar or ostac or clastoban or lytos or clasteon or climaclod or codeosden or clody or difosfonal or dolkin or moticlod or niklod or ositten or osteonorm or osteostab or mebonat (85)
S27 alendronate or "alendronic acid" or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or oseotenk or osteofene or osteonate or phostarac or regeness or silidril or tillos or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossofax or osteofar or osteofem or osteoral or osteotrat or recalfe or terost or aldrox or fosval or holaden or leodrin or oseotenk or osteofene or osteonate or osfalen or splset or bifosa or osteofe or fosalan or maxibone or adronat or lendrons or drolon or genalen or onclast or drovitan or landrofen or sinfrac or aldrofac or aliott or denfos or fixopan or genalmor or osteoedur or porosalaendronate or "alendronic acid" or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or osteotenk or osteofene or osteonate or phostarac or regeness or silidril or tillos or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossmax or ostenan or osteofar or osteofem or osteoral or osteotrat or recalfe or terost or aldrox or...
or fosval or holadren or leodrin or oseotal or osteofem or osteoan or alefos or fosalen or ostalert or bifosa or osteofos or fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or landrolen or sinfract or aldronac or aliot or denfos or fixopan or genalmen or osteodur or porosal (765)
S26 bisphosphonate* or diphosphonate* (3182)
S25 (MH "Diphosphonates") or (MH "alendronate") (3160)
S24 S22 or S23 (10,599)
S23 osteoporo* (10,505)
S22 (MH "Osteoporosis") (8995)
S21 S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S20 (772)
S20 S18 and S19 (164)
S19 (MH "Bone and Bones") (2891)
S18 (MH "Biological Markers") (12,859)
S17 "bone marker*" (108)
S16 ("biochemical marker*" or biomarker* or "biological marker*") n2 bone* (224)
S15 "bone metabolic marker*" (4)
S14 "bone turnover marker*" (129)
S13 c-telopeptide* n3 collagen (44)
S12 n-telopeptide* n3 collagen (42)
S11 telopeptide* n3 collagen (168)
S10 crosslaps (16)
S9 (CTX or NTX) and (bone or bones or biomarker* or biological marker*) (143)
S8 "bone source alkaline phosphatase*" (0)
S7 "bone alkaline phosphatase*" (57)
S6 "bone specific alkaline phosphatase*" (101)
S5 (BSAP or BALP or BAP) and (bone or bones or biomarker* or biological marker*) (61)
S4 collagen n3 propeptide (37)
S3 procollagen n3 peptide (18)
S2 procollagen n3 propeptide (45)
S1 P1NP or PINP (25)

Key
MH = indexing term (CINAHL heading)
+ = exploded CINAHL heading
* = truncation
? = embedded truncation
* * = phrase search
n2 = terms within one word of each other (any order)
n3 = terms within two words of each other (any order)

ClinicalTrials.gov
URL: www.clinicaltrials.gov.
Date searched: 13 March 2012.
Records found: 98.
Search terms
osteoporosis AND ("bone markers" OR "bone turnover markers")

Key
* " = phrase search

The Cochrane Library

Issue 2 of 12 February 2012.

Date searched: 12 March 2012.

Records found:

- Cochrane Database of Systematic Reviews 30
- Database of Abstracts of Reviews of Effects (DARE) 5
- Cochrane Central Register of Controlled Trials 496
- NHS Economic Evaluation Database 1
- Health Technology Assessment Database 4

Search terms
#1 P1NP or PINP (133)
#2 procollagen near/3 propeptide (68)
#3 procollagen near/3 peptide (140)
#4 collagen near/3 propeptide (26)
#5 (BSAP or BALP or BAP) and (bone or bones or biomarker* or biological marker*) (189)
#6 "bone specific alkaline phosphatase*" (385)
#7 "bone alkaline phosphatase*" (189)
#8 "bone source alkaline phosphatase*" (0)
#9 (CTX or NTX) and (bone or bones or biomarker* or biological marker*) (464)
#10 crosslaps (55)
#11 telopeptide* near/3 collagen (152)
#12 "n-telopeptide*" near/3 collagen (37)
#13 "c-telopeptide*" near/3 collagen (45)
#14 "bone turnover marker*" (23)
#15 "bone metabolic marker*" (3)
#16 ("biochemical marker*" or biomarker* or "biological marker*") near/2 bone* (108)
#17 "bone marker*" (55)
#18 MeSH descriptor Biological Markers, this term only (6199)
#19 MeSH descriptor Bone and Bones, this term only (1156)
#20 (#18 AND #19) (258)
#21 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #20) (1370)
#22 MeSH descriptor Osteoporosis explode all trees (2750)
#23 osteoporosen* (4841)
#24 (#22 OR #23) (4841)
#25 MeSH descriptor Diposphonates, this term only (738)
#26 MeSH descriptor Alendronate, this term only (498)
#27 MeSH descriptor Clodronic Acid, this term only (166)
#28 MeSH descriptor Etidronic Acid, this term only (370)
#29 bisphosphonate* or diphosphonate* (1336)

DOI: 10.3310/hta18110

© Queen’s Printer and Controller of HMSO 2014. This work was produced by Burch et al. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.
#30 alendronate or "alendronic acid" or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or osteofen or osteone or phostarac or regenesis or silidral or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossomax or ostalen or osteofar or osteoform or osteoral (775)
#31 osteotrat or recalfe or terost or aldrox or fosval or holadren or leodrin or oseotal or osteofem or osteosan or alenox or fosalen or ostalert or bifosa or osteofos or fosalan or maxibone or adronat or alendros or dronal or marvil or maxtral or osteofen or osteone or phostarac or regenesis or silidral or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossomax or osteonal (51)
#33 osteofar or osteoform or osteoral or osteotrat or recalfe or terost or aldrox or fosval or holadren (2)
#34 leodrin or oseotal or osteofem or osteosan or alenox or fosalen or ostalert or bifosa or osteofos (0)
#35 fosalan or maxibone or adronat or alendros or dronal or marvil or maxtral or osteofen or osteone or phostarac or regenesis or silidral or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossomax or osteonal (2)
#36 clodronate or "clodronic acid" or bonefos or loron or clodron or lodronat or ascredar or ostac or clastoban or lytos or clasteon or cladoclor or clostalien or clodeosten or clODY or clodin or clotir or clody or clonid or coddone or coddon or clonan or clonate or clonate-os or clodronate or clotro or clotro (535)
#38 "ibandronic acid" or ibandronate or boniva or bondronat or bonviva or bandrobon or bonat (177)
#39 pamidronate or "pamidronic acid" or aredria or "ADP sodium" or aminomux or pandosa or pandrat or pamisol or pamitol or osteopam or aredonet or pamidron or pamidran or linoten or xinsidona (422)
#40 risenedronate or "risenedronic acid" or actonel or ductonar or ribastamin or ridron or risedon or optinate or risedronate or benet or acril (350)
#41 "zoledronic acid" or zoledronate or zometa or zomerla or aclast or reclast or zoldria (373)
#42 "tiludronic acid" or tiludronate or skelid (29)
#43 "neridronic acid" or neridronate or nerixia (23)
#44 "olpadronic acid" or olpadronate (12)
#45 "cimadronic acid" or cimadronate (0)
#46 "piridronic acid" or piridronate (0)
#47 "icandronic acid" or icandronate or bisphonal (0)
#48 "minodronic acid" or minodronate (2)
#49 raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax or extroact or ralista or celvista (578)
#50 MeSH descriptor Raloxifene, this term only (373)
#51 "strontium ranelate" or protelos (59)
#52 denosumab or prolia or xgeva (64)
#53 MeSH descriptor Teriparatide, this term only (126)
#54 treatment* or treat or treated or treats (330,609)
#55 (#25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40) (2500)
#56 (#41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54) (330,912)
#57 (#55 OR #56) (331,447)
#58 (#21 AND #24 AND #57) (536)
Key
MeSH descriptor = indexing term (MeSH heading)
* = truncation
" " = phrase search
near/2 = terms within two words of each other (any order)
near/3 = terms within three words of each other (any order)

Conference Proceedings Citation Index – Science (Web of Science)

Date range: 1990–date.

Date searched: 12 March 2012.

Records found: 197.

Search terms
# 45 #44 AND #19 AND #18
# 44 #43 OR #42 OR #41 OR #40 OR #39 OR #38 OR #37 OR #36 OR #35 OR #34 OR #33 OR #32
OR #31 OR #30 OR #29 OR #28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20
# 43 Topic=(treatment* or treat or treated or treats)
# 42 Topic=(denosumab or prolia or xgeva)
# 41 Topic=("strontium ranelate" or protelos)
# 40 Topic=(raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax
or estroact or ralista or celvista)
# 39 Topic=("minodronic acid" or minodronate)
# 38 Topic=("icandronic acid" or icandronate or bisphonal)
# 37 Topic=("piridronic acid" or piridronate)
# 36 Topic=("cimadronic acid" or cimadronate)
# 35 Topic=("olpadronic acid" or olpadronate)
# 34 Topic=("neridronic acid" or neridronate or nerixia)
# 33 Topic=("tiludronic acid" or tiludronate or skelid)
# 32 Topic=("zoledronic acid" or zoledronate or zometa or zomera or aclasta or reclar or zoldria)
# 31 Topic=(risedronate or *risedronic acid* or actionel or ductonar or ribastamin or ridron or risedon or
optinate or risofos or benet or acrel)
# 30 Topic=(pamidronate or "pamidronic acid" or aredia or "ADP sodium" or aminomux or pamdosa or
pandrat or pamisol or pamitor or osteopam or aredronet or pamidria or pamifos or pamidran or linoten
or xinsidon)
# 29 Topic=("ibandronic acid" or ibandronate or boniva or bononar or bonviva or bandrobon or bonat)
# 28 Topic=(etidronate or "etidronic acid" or EHDP or didronel or difosfen or detidron or osteidronel or
bonemass or osteotop or didronate or diphos or anfozan or biotredine or dralen or etidron or etiplus or
feminoflex or maxibral or oflocin or oslo or osteron or osteodrug or osteotcon or ostogene or somaflex
or sterodome or sviroxit or tilferan or dronate-os or etidrate or osteum)
# 27 Topic=(clodronate or "clodronic acid" or bonefors or leron or clodron or lodronat or ascendar or ostac
or clastonan or lytos or clasteon or climaclod or clodeosten or clody or difosfonal or dolkin or moticlod or
niklod or ositen or osteonorm or osteostab or mebonat)
# 26 Topic=(fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or
landrolen or sinfract or aldronac or aliot or denfos or fixopan or genalmen or osteodur or porosal)
# 25 Topic=(leodrin or oseotal or osteofem or osteosan or alefos or fosalen or ostalert or bifosa
or osteofos)
# 24 Topic=(osteofar or osteoform or osteoral or osteotrat or recalfe or terost or aldrox or fosval
or holadren)
# 23 Topic=(actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or oseotenk or osteofene or osteonate or phostarac or regenesis or silidral or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossomax or ostenan)
# 22 Topic=(osteotrat or recalfe or terost or aldrox or fosval or holadren or leodrin or osteofem or osteoan or alefos or fosalen or ostalert or bifosa or osteofos or fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or landrolen or sinfract or aldronac or aliot or denfos or fixopan or genalmen or osteofar or osteoform or osteoral)
# 21 Topic=(alendronate or "alendronic acid" or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or oseotenk or osteofene or osteonate or phostarac or regenesis or silidral or tilios or bonalen or cleveron or minusorb or norvic or osdron or ossomax or ostenan or osteofar or osteoform or osteoral)
# 20 Topic=(bisphosphonate* or diphosphonate*)
# 19 Topic=(osteoporos*)
# 18 #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
# 17 Topic="bone marker*"
# 16 Topic="("biochemical marker*" or biomarker* or "biological marker*") near/2 bone*)
# 15 Topic="bone metabolic marker*"
# 14 Topic="bone turnover marker*"
# 13 Topic="c-telopeptide*" near/3 collagen)
# 12 Topic="n-telopeptide*" near/3 collagen)
# 11 Topic="telopeptide* near/3 collagen)
# 10 Topic=crosslaps)
# 9 Topic=((CTX or NTX) and (bone or bones or biomarker* or biological marker*))
# 8 Topic="bone source alkaline phosphatase*)"
# 7 Topic="bone alkaline phosphatase*)"
# 6 Topic="bone specific alkaline phosphatase*)"
# 5 Topic=((BSAP or BALP or BAP) and (bone or bones or biomarker* or biological marker*))
# 4 Topic=(collagen near/3 propeptide)
# 3 Topic=(procollagen near/3 peptide)
# 2 Topic=(procollagen near/3 propeptide)
# 1 Topic=(P1NP or PINP)

Limits: Lemmatization – OFF

Key
Topic = terms in Title, Abstract, Author Keywords and Keywords Plus fields
* = truncation
? = embedded truncation
** = phrase search
near/2 = terms within one word of each other (any order)
near/3 = terms within two words of each other (any order)

Controlled-Trials.com


Date searched: 13 March 2012.

Records found: 99.
Search terms
osteoporosis AND ("bone markers" OR "bone turnover markers")

Key
" " = phrase search

EconLit (Ovid)

Date range: 1961–February 2012.
Date searched: 7 March 2012.
Records found: none.

Search terms
1. (P1NP or P1NP).ti,ab. (0)
2. (procollagen adj3 propeptide).ti,ab. (0)
3. (procollagen adj3 peptide).ti,ab. (0)
4. (collagen adj3 propeptide).ti,ab. (0)
5. (BSAP or BALP or BAP).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (0)
6. bone specific alkaline phosphatase$.ti,ab. (0)
7. bone alkaline phosphatase$.ti,ab. (0)
8. bone source alkaline phosphatase$.ti,ab. (0)
9. (CTX or NTX).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (0)
10. crosslaps.ti,ab. (0)
11. (telopeptide$ adj3 collagen).ti,ab. (0)
12. (n-telopeptide$ adj3 collagen).ti,ab. (0)
13. (c-telopeptide$ adj3 collagen).ti,ab. (0)
14. bone turnover marker$.ti,ab. (0)
15. bone metabolic marker$.ti,ab. (0)
16. ((biochemical marker$ or biomarker$ or biological marker$) adj2 bone$).ti,ab. (0)
17. bone marker$.ti,ab. (0)
18. or/1–17 (0)
19. osteoporos$.ti,ab. (24)
20. (bisphosphonate$ or diphosphonate$).af. (2)
21. (alendronate or alendronic acid or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or osenotenk or osteofene or osteonate or phostarac or regenesis or siildiral or tilios or bonalen or cleveron or endrnax or minusorb or norvic or osdron or ossomax or ostenan or osteofar or osteoform or osteoral or osteotrat or recaffe or terost or aldrox or fosval or holadren or leodrin or oseotal or osteosan or alefos or fosalen or ostalert or bifosa or osteofos or fosalan or maxibone or adronat or adrenal or drental or dromal or genalen or onclast or drovitam or landronal or sinfract or aldranon or alliot or denfoss or fixopan or genmalen or osteodur or porosal).af. (5)
22. (clodronate or clodronic acid or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvil or maxtral or osenotenk or osteofene or osteonate or phostarac or regenesis or siildiral or tilios or bonalen or cleveron or endrnax or minusorb or norvic or osdron or ossomax or ostenan or osteofar or osteoform or osteoral or osteotrat or recaffe or terost or aldrox or fosval or holadren or leodrin or oseotal or osteosan or alefos or fosalen or ostalert or bifosa or osteofos or fosalan or maxibone or adronat or adrenal or drental or dromal or genalen or onclast or drovitam or landronal or sinfract or aldranon or alliot or denfoss or fixopan or genmalen or osteodur or porosal).af. (0)
23. (etidronate or etidronic acid or EHDP or didronel or disosfen or detidron or osteodidronel or bonemass or osteotop or diphos or anfrozan or biotredine or dralen or etidron or etiplus or feminoflex or maxbral or oflocin or osfo or ostreden or osteodrug or osteotin or osteosteg or somaflex or sterodome or sviroxit or tiiferan or dronate-os or etidrate or osteum).af. (0)
24. (ibandronic acid or ibandronate or boniva or bondronat or bonviva or bandrobom or bonat).af. (0)
APPENDIX 1

25. (pamidronate or pamidronic acid or aredia or ADP sodium or aminomux or pamdosa or pandrat or pamisol or pamitor or osteopam or aredronet or pamidria or pamifos or pamidran or linoten or xinsidona).af. (0)
26. (risedronate or risedronic acid or actonel or ductonar or ribastamin or ridron or risedon or optinate or risofos or benet or acrel).af. (7)
27. (zoledronic acid or zoledronate or zometa or zomera or aclasta or reclast or zoldria).af. (1)
28. (tiludronic acid or tiludronate or skelid).af. (1)
29. (neridronic acid or neridronate or nerixia).af. (0)
30. (olpadronic acid or olpadronate).af. (0)
31. (cimadronic acid or cimadronate).af. (0)
32. (minodronic acid or minodronate).af. (0)
33. (raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax or estroact or ralista or celvista).af. (2)
34. (strontium ranelate or protelos).af. (0)
35. (treatment$ or treat or treated or treats).ti,ab. (15,352)
40. or/20-39 (15363)
41. 18 and 19 and 40 (0)

Key
$ = truncation
.ti,ab. = terms in either title or abstract fields
.af. = terms in all fields
adj = terms adjacent to each other (same order)
adj2 = terms within two words of each other (any order)
adj3 = terms within three words of each other (any order)

EMBASE (Ovid)

Date range: 1974–6 March 2012.

Date searched: 7 March 2012.

Records found: 2495.

Search terms
1. (P1NP or PINP).ti,ab. (908)
2. (procollagen adj3 propeptide).ti,ab. (511)
3. (procollagen adj3 peptide).ti,ab. (663)
4. (collagen adj3 propeptide).ti,ab. (225)
5. (BSAP or BALP or BAP).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (1440)
6. bone specific alkaline phosphatase$.ti,ab. (1388)
7. bone alkaline phosphatase$.ti,ab. (1421)
8. bone source alkaline phosphatase$.ti,ab. (1)
9. (CTX or NTX).ti,ab. and (bone or bones or biomarker$ or biological marker$).af. (2846)
10. crosslaps.ti,ab. (429)
11. (telopeptide$ adj3 collagen).ti,ab. (839)
12. (n-telopeptide$ adj3 collagen).ti,ab. (251)
13. (c-telopeptide$ adj3 collagen).ti,ab. (228)
14. bone turnover marker$.ti,ab. (1780)
15. bone metabolic marker$.ti,ab. (232)
16. (marker/ or biochemical marker/ or biological marker/ or disease marker/) and (bone/ or bone turnover/) (3965)
17. ((biochemical marker$ or biomarker$ or biological marker$) adj2 bone$).ti,ab. (2782)
18. bone marker$.ti,ab. (1800)
19. or/1-18 (12,593)
20. exp osteoporosis/ (75,240)
21. osteoporos$.ti,ab. (58,711)
22. 20 or 21 (88,169)
23. bisphosphonic acid derivative/ or alendronic acid/ or clodronic acid/ or etidronic acid/ (29,528)
24. (bisphosphonate$ or diphosphonate$).af. (16,503)
25. (alendronate or alendronic acid or fosamax or actimax or alenato or arendal or brek or elandur or findeclin or lafedam or leandroal or marvil or maxtral or oseotenik or osteofene or osteonate or phostarac or regenesis or silidral or tilos or bonalen or cleveron or endronax or minusorb or norvic or osdran or ossomax or ostenan or osteofar or osteoform or osteoral or osteotrat or recalfe or terost or aldrox or fosval or holadren or leodrin or oseotal or osteofem or osteosan or alefos or fosalen or osteostal or bifosa or osteofos or fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or landrolley or sinfract or aldonac or ailet or denfos or fixopan or genalmen or osteodur or porosal).af. (14,288)
26. (clodronate or clodronic acid or bonefos or loron or clodron or lodronat or ascnedir or ostac or clastoban or lytos or clasteon or climaxod or cloak or dioxifon or dolkin or motiolod or niklod or ossiten or osseotrin or osteostab or mebonat).af. (4623)
27. (etidronate or etidronic acid or EHDP or didronel or diphos or etidron or osteodine or bonemass or osteopt or didronate or diphos or anfozan or biotredine or dralen or etidron or etiplus or feminoflex or maxbral or offocin or osfo or osedron or osteodrug or osteon or somaflex or sterdome or sviroxit or tilferan or dronate-os or etidrate or osteum).af. (7748)
28. (ibandronic acid or ibandronate or boniva or bondronat or bonivina or bandrobon or bonat).af. (3253)
29. (pamidronate or pamidronic acid or aredia or ADP sodium or aminomux or pamdosa or pandrat or pamisol or pamitum or osteopam or aredronet or pamidria or pamifos or pamidran or linoten or xisidona).af. (7680)
30. (risedronate or risedronic acid or actonel or ductoran or ribastamin or ridron or risedon or optinate or risosofos or benet or acrel).af. (6164)
31. (zoledronic acid or zoledronate or zometa or zomora or aclast or reclast or zoldria).af. (7257)
32. (tiludronic acid or tiludronate or skelid).af. (744)
33. (neridronic acid or neridronate or nerixia).af. (266)
34. (olpadronic acid or olpadronate).af. (242)
35. (cimadronic acid or cimadrone).af. (9)
36. (piridronic acid or piridronate).af. (0)
37. (candronic acid or icandronate or bisphonal).af. (2)
38. (minodronic acid or minodronate).af. (182)
39. (raloxifene or biofem or ciclotran or evista or ketidin or loxifen or optruma or bonmax or estroact or ralista or celvista).af. (8201)
40. raloxifene/ (7977)
41. (strontium ranelate or protelos).af. (1325)
42. "parathyroid hormone[1-34]"/ (3339)
43. (teriparatide or forteo or forsteo).ti,ab. (913)
44. (denosumab or prolia or xgeva).af. (1737)
45. (treatment$ or treat or treated or treats).ti,ab. (3,950,972)
46. dt.fs. (2,648,356)
47. or/23-46 (5,432,565)
48. 19 and 22 and 47 (2670)
49. exp animal/ or exp nonhuman/ (5,469,091)
50. exp human/ (13070114)
51. 49 not (49 and 50) (4,361,154)
52. 48 not 5 (2495)

Key
/ = indexing term (EMTREE heading)
exp = exploded EMTREE heading
$ = truncation
.ti,ab. = terms in either title or abstract fields
.af. = terms in all fields
.fs. = floating subheading – searches all EMTREE terms which use that subheading
adj = terms adjacent to each other (same order)
adj2 = terms within two words of each other (same order)
adj3 = terms within three words of each other (any order)

Paid Clinical Trials
URL: www.paidclinicaltrials.org.
Date searched: 1 May 2012.
Records found: none.

Search terms
Browsed: osteoporosis

Science Citation Index Expanded (Web of Science)
Date range: 1899–date.
Date searched: 12 March 2012.
Records found: 2085.

Search terms
# 45 #44 AND #19 AND #18
# 44 #43 OR #42 OR #41 OR #40 OR #39 OR #38 OR #37 OR #36 OR #35 OR #34 OR #33 OR #32 OR #31 OR #30 OR #29 OR #28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20
# 43 Topic=(treatment* or treat or treated or treats)
# 42 Topic=(denosumab or prolia or xgeva)
# 41 Topic=(*strontium ranelate* or protelos)
# 40 Topic=(raloxifene or biofem or ciclotran or evista or ketidin or loxifen or raxeto or optruma or bonmax or estroact or ralista or celvista)
# 39 Topic=(*minodronic acid* or minodronate)
# 38 Topic=(*icandronic acid* or icandronate or bisphonal)
# 37 Topic=(*piridronic acid* or piridronate)
# 36 Topic=(*cimadronic acid* or cimadronate)
# 35 Topic=(*olpadronic acid* or olpadronate)
# 34 Topic=(*neridronic acid* or neridronate or nerixia)
# 33 Topic=(*tiludronic acid* or tiludronate or skelid)
# 32 Topic=(*zoledronic acid* or zoledronate or zometa or zomera or aclasta or reclast or zoldria)
# 31 Topic=(risedronate or "risedronic acid" or actonel or ductionar or ribastamin or ridron or risedon or optinate or risofos or benet or acrel)
# 30 Topic=(pamidronate or "pamidronic acid" or aredia or "ADP sodium" or aminomux or pandos or pandrat or pamisol or pamitor or osteopam or arerdonet or pamidia or pamifos or pamidran or linoten or xinsidona)
# 29 Topic=("ibandronic acid" or ibandronate or boniva or bonilux or bonviva or bandrobon or bonat)
# 28 Topic=(etidronate or "etidronic acid" or EHDP or didronel or didfosfen or detidron or osteidronel or bonemass or osteotop or idronate or diphos or anfozan or biotredine or dralen or etidron or etiplus or feminoflex or maxbral or ofilocin or osfo or osteron or osteodrug or osteoton or osteogene or somaflex or sterodome or siviroix or tilferan or dromate-os or etidrate or osteum)
# 27 Topic=(clodronate or "clodronic acid" or bonefos or loron or clodron or lodronat or ascredar or ostac or clastoban or lytos or clasto or climaclod or clodeosten or clody or difosonal or dolkin or moticlod or niklod or ossiten or osteonorm or osteostab or mebonat)
# 26 Topic=(fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or landrolien or sinfract or alidronac or alefon or osfalon or ostaliert or bifosa or osteofos)
# 25 Topic=(leodrin or oseotal or osteofem or osteosan or alefos or fosalen or ostalert or bifosa or osteofos)
# 24 Topic=(osteofar or osteoform or osteoral or osteotrat or calcalf or terost or aldros or fosval or holaden)
# 23 Topic=(actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvila or maxtral or oseotenk or osteofene or osteonate or phostarac or regenesis or silidril or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossmax or ostenan)
# 22 Topic=(osteocalc or osteocalf or terost or aldros or fosval or osteron or holaden or leodrin or oseotal or osteofem or osteosan or alefos or fosalen or ostaliert or bifosa or osteofos or fosalan or maxibone or adronat or alendros or dronal or genalen or onclast or drovitan or landrolien or sinfract or alidronac or alefon or osfalon or ostaliert or bifosa or osteofos)
# 21 Topic=(alendronate or "alendronic acid" or fosamax or actimax or alenato or arendal or berlex or brek or elandur or findeclin or lafedam or lendronal or marvila or maxtral or oseotenk or osteofene or osteonate or phostarac or regenesis or silidril or tilios or bonalen or cleveron or endronax or minusorb or norvic or osdron or ossmax or ostenan or osteofar or osteoform or osteoral)
# 20 Topic=(bisphosphonate* or diphosphonate*)
# 19 Topic=(osteopor*)
# 18 #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
# 17 Topic=("bone marker**)
# 16 Topic=(("biochemical marker*" or biomarker* or *biological marker*) near/2 bone*)
# 15 Topic=("bone metabolic marker"**)
# 14 Topic=("bone turnover marker"**)
# 13 Topic=("c-telopeptide** near/3 collagen)
# 12 Topic=("n-telopeptide"** near/3 collagen)
# 11 Topic=(telopeptide* near/3 collagen)
# 10 Topic=(crosslaps)
# 9 Topic=((CTX or NTX) and (bone or bones or biomarker* or biological marker*))
# 8 Topic=("bone source alkaline phosphatase**")
# 7 Topic=("bone alkaline phosphatase**")
# 6 Topic=("bone specific alkaline phosphatase**")
# 5 Topic=((BSAP or BALP or BAP) and (bone or bones or biomarker* or biological marker*))
# 4 Topic=(collagen near/3 propeptide)
# 3 Topic=(procollagen near/3 peptide)
# 2 Topic=(procollagen near/3 propeptide)
# 1 Topic=(P1NP or PINP)
Key
Topic = terms in Title, Abstract, Author Keywords and Keywords Plus fields
* = truncation
? = embedded truncation
* * = phrase search
near/2 = terms within one word of each other (any order)
near/3 = terms within two words of each other (any order)

IDEAS database
URL: http://ideas.repec.org.
Date searched: 15 May 2012
Records found: nine [after hand-sifting for relevance].

Search terms
treatment AND adherence [title]
treatment AND monitoring [title]
treatment AND compliance [title]
osteoporosis AND adherence [title]
osteoporosis AND monitoring [title]
osteoporosis AND compliance [title]

Health Economic Evaluations Database
URL: www.cochrane.org/intranet/resources-databases/health-economics-evaluation-database-heed.

Search 1
Date range: all to date.
Date searched: 13 March 2012.
Records found: six.

Search terms
All data: osteoporosis

AND

All data: marker*
Search 2
Date range: all to date.

Date searched: 15 May 2012.

Records found: 49.

Search terms
All data: osteoporosis

AND

All data: monitor* or adher* or comply or compliance or complies or complied

Search 3
Date range: all to date.

Date searched: 15 May 2012.

Records found: seven.

Search terms
All data: ‘treatment monitoring’

OR

All data: ‘monitoring treatment’

OR

All data: ‘monitor treatment’

Key
* = truncation
’ ’ = phrase searching

NHS Economic Evaluation Database

The Cochrane Library, Issue 4 of 12 April 2012.

Date searched: 15 May 2012.

Records found: 79.
Search terms
#1 MeSH descriptor Osteoporosis explode all trees 2785
#2 osteoporo* 4914
#3 MeSH descriptor Patient Acceptance of Health Care explode all trees 16,473
#4 monitor* or adher* or comply or compliance or complies or complied 56,037
#5 (#1 OR #2) 4914
#6 (#3 OR #4) 64,274
#7 (#5 AND #6) 930
#8 treatment NEAR/1 monitor* 217
#9 (#7 OR #8) 79 [NHS EED database only]

Key
MeSH descriptor = indexing term (MeSH heading)
* = truncation
near/1 = terms within one word of each other (any order)
Appendix 2  Results of and guidelines for the quality assessment
### Randomised controlled trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Sample</th>
<th>Randomisation</th>
<th>Allocation</th>
<th>Baseline comparability</th>
<th>Blinding</th>
<th>Study aim</th>
<th>Intervention details</th>
<th>Population representative</th>
<th>Control group selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delmas (2007)</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kung (2009)</td>
<td>UC</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roche (2009)</td>
<td>UC</td>
<td>Y</td>
<td>UC</td>
<td>UC</td>
<td>UC</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roche (2009)</td>
<td>UC</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>UC</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roche (2007)</td>
<td>UC</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>UC</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N, no; P, partially (modified ITT population); UC, unclear; Y, yes.

### Non-randomised controlled trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Sample</th>
<th>Randomisation</th>
<th>Allocation</th>
<th>Baseline comparability</th>
<th>Blinding</th>
<th>Study aim</th>
<th>Intervention details</th>
<th>Population representative</th>
<th>Control group selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller (2008)</td>
<td>Controlled cohort: consecutive recruitment</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ishijima (2009)</td>
<td>Uncontrolled cohort: consecutive recruitment</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majima (2008)</td>
<td>Uncontrolled cohort: consecutive recruitment</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shiraki (2011)</td>
<td>Uncontrolled cohort: consecutive recruitment</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imai (2009)</td>
<td>Uncontrolled cohort: non-consecutive recruitment</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tsujimoto (2011)</td>
<td>Derived cohort: post hoc analysis of RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heaney (2011)</td>
<td>Derived cohort: post hoc analysis of RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>P</td>
<td>UC</td>
<td>N</td>
<td>&lt; 1</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Some analyses</td>
<td>UC</td>
<td>N</td>
<td>&lt; 1</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>UC</td>
<td>N</td>
<td>N</td>
<td>&lt; 1</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>N</td>
<td>N</td>
<td>≥ 1</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>Y</td>
<td>N</td>
<td>&lt; 1</td>
<td>Low</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>N</td>
<td>P</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>&lt; 1</td>
<td>Y</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>&lt; 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>UC</td>
<td>N</td>
<td>&lt; 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>UC</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>&lt; 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>UC</td>
<td>NA</td>
<td>NA</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>NA</td>
<td>NA</td>
<td>Low</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>Y</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Design</td>
<td>Population</td>
<td>Intervention</td>
<td>Aim</td>
<td>Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>--------</td>
<td>------------</td>
<td>--------------</td>
<td>-----</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastell (2011)</td>
<td>63</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bruyere (2010)</td>
<td>141</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reginster (2004)</td>
<td>124</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitatani (2003)</td>
<td>156</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2005)</td>
<td>140</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bjarnason (2001)</td>
<td>151</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane (2000)</td>
<td>159</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dobnig (2005)</td>
<td>152</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bauer (2004)</td>
<td>153</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dobnig (2006)</td>
<td>154</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clowes (2003)</td>
<td>47</td>
<td>Derived cohort: treatment arm(s) from RCT</td>
<td>UC</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blumsohn (2011)</td>
<td>155</td>
<td>Derived cohort: RCT non-randomised extension</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim (2005)</td>
<td>156</td>
<td>Uncontrolled cohort: UC</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reyes-Garcia (2010)</td>
<td>157</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masaryk (2002)</td>
<td>158</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyd (1998)</td>
<td>159</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyd (1999)</td>
<td>160</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2005)</td>
<td>161</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2004)</td>
<td>162</td>
<td>Uncontrolled cohort: UC</td>
<td>Y</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armstrong (2007)</td>
<td>143</td>
<td>Uncontrolled cohort: UC</td>
<td>UC</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stepan (2008)</td>
<td>150</td>
<td>Uncontrolled cohort: UC</td>
<td>UC</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro-Alvarez (2010)</td>
<td>154</td>
<td>Uncontrolled cohort: UC</td>
<td>UC</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddiqi (2010)</td>
<td>155</td>
<td>Uncontrolled cohort: UC</td>
<td>N</td>
<td>NA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N, no; NA, not applicable; P, partially (modified ITT population); UC, unclear; Y, yes.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>Y</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>UC</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>UC</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>UC</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≤ 1</td>
<td>N</td>
<td>UC</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>UC</td>
<td>UC</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>≥ 1</td>
<td>N</td>
<td>N</td>
<td>Low</td>
</tr>
</tbody>
</table>
Guidelines for completing the quality assessment

1. Patient selection

Randomised – method appropriate: random numbers computer generated or number table, controlled by external source (pharmacy, biochemistry laboratory), or similar

Randomised – method not reported: date of birth, day of recruitment or similar

Randomised – no details: states patients were randomised but did not report method used

Consecutive recruitment: non-randomised study: all patients recruited consecutively

Non-consecutive recruitment: non-randomised study: selective recruiting

Post hoc analysis of prior RCT: results of patients recruited prospectively into a RCT were reanalysed

Post hoc analysis of prior cohort: results of patients recruited prospectively into a cohort study were reanalysed

Random selection (derived cohort): patients were recruited prospectively into a RCT comparing an antiresorptive drug of interest with placebo or an alternative drug; those receiving a drug of interest were treated in the review as a prospective cohort

Patients were selected for an open-label extension of a RCT

Unclear: ‘patient selection’ mentioned but no details as to method used

Not reported: no mention of a patient selection process, just number included in the study

2. Population representative

Yes: the population in the study is representative of that expected in clinical practice (i.e. a heterogeneous population of patients with osteoporosis or a study recruiting unselected post-menopausal women with osteoporosis); population details includes the drug regimen

No: the population in the study is not representative of that expected in clinical practice (i.e. it is in a subgroup of patients), a selected population of post-menopausal women, or it is a study in a heterogeneous population that excludes specific subgroups

Unclear: there was insufficient information to determine whether or not the population in the study is representative of that expected in clinical practice

3. Allocation concealment

Method appropriate: sequentially numbered sealed opaque envelopes or containers, controlled by external source (pharmacy, biochemistry laboratory), or similar

Method inappropriate: the method used was not one of those stated above, for example date of birth, day of recruitment

Unclear: details regarding allocation concealment were not reported

Not applicable: single-arm study or data from a RCT used as a derived cohort

4. Control group selection appropriate

Yes: part of an adequate randomisation/allocation concealment in RCTs, or drawn from same population at the same time in observational studies

No: drawn from a different population or time (e.g. historical control)

Unclear: details regarding the selection of controls were not reported

Not applicable: single-arm study or data from a RCT used as a derived cohort

5. Baseline comparability

Yes: baseline characteristics were similar across groups

No: baseline characteristics were not similar across groups

Unclear: insufficient details provided to determine similarity across groups at baseline

Not applicable: single-arm study or data used as a derived cohort from a RCT
### 6a. Blinding

Unblinded: patients randomised but no blinding or controlled cohort study

Single-blind: reports being single-blind

Double-blind: reports being double-blind

Triple-blind: reports being triple-blind

Open-label extension: patients followed in extension of RCT where there was no continued blinding

Unclear: blinding status was not reported

Not applicable: single-arm study

### 6b. Who was blinded

Patients: clear statement that patients were blinded

Carers/investigators: clear statement that carers/investigators were blinded

Outcome assessors: clear statement that outcome assessors were blinded

Unclear: blinding was indicated but who was blinded was not specified

Not applicable: study was unblended or a single-arm study

### 7. Clear descriptions of study aim

Yes: aim of the study clearly stated

No: no clear statement as to the aim of the study

### 8. Clear descriptions of intervention details

Yes: details of the bone turnover marker and other tests clearly stated to allow replication

No: insufficient details of the bone turnover marker and other tests to allow replication

### 9. Clear descriptions of population details

Yes: details of the population clearly stated to allow an assessment of the representativeness of the population

No: population not clearly described

### 10. Clear descriptions of main findings

Yes: main study findings clearly described

No: no clear description of the main findings

### 11. Loss to follow-up: characteristics described

Yes: the characteristics of those lost to follow-up sufficiently described to allow comparison with those who remained in the study, or a comparison was made between the two groups by the study authors and a statement made such as whether or not loss was random

No: characteristics of those who dropped out were not reported and no comparison with those who remained in the study was made

### 12. Loss to follow-up: reasons given

Yes: the reasons patients were lost to follow-up were given or there were no losses to follow-up

No: the reasons patients were lost to follow-up were not reported

### 13. Loss to follow-up: taken into account in analysis

Yes: patients lost to follow-up were included in the analysis or there were no losses to follow-up

No: a completer analysis was conducted

Unclear: it was unclear whether or not patients lost to follow-up were included in the analysis
14. Loss to follow-up: imputation methods appropriate

Yes: an ITT analysis, last observation or baseline carried forward, or best-worst-case scenario analyses conducted (or there were no losses to follow-up)

No: other methods of imputation used, or data were not imputed and used in the analysis

Unclear: method of imputation not reported

15. Measure of variance reported round estimates

Yes: a measure of variance was reported around the point estimate

No: a measure of variance was not reported around the point estimate; this not inappropriate for a correlation coefficient

16. Adverse events reported

Yes: adverse events associated with the BMs were reported

No: adverse events associated with the BMs were not reported

Not applicable: the study was a derived cohort where BMs were an outcome measure not an intervention

17. Duration of follow-up

< 1 year: minimum follow-up was less than 1 year

≥ 1 year: follow-up was at least 1 year in all patients

Unclear: the duration of follow-up was not reported

18. Confounders: clearly described

Yes: potentially confounding factors described

No: potentially confounding factors not described

Not applicable: RCT with appropriate methods of randomisation

19. Confounders: all important confounders adjusted for

Yes: all important confounding factors (age, gender, prior fracture, baseline BMD, BMI) were adjusted for in the analysis

No: adjustments were made but not for all those considered important, or there was not adjustment for confounding factors

Unclear: it was not clear whether adjustments were made or, if they were, the variables adjusted for were not reported

Not applicable: RCT with appropriate methods of randomisation and ITT analysis

Overall quality:

High: the study is not subject to bias, or the bias/limitations of the study will not impact on the reliability of the results

**RCTs:**
- Appropriate randomisation and allocation concealment methods used; patients, carers, investigators and outcome assessors blinded; ITT results reported using established imputation methods; minimum of 1-year follow-up in all patients

**Cohort studies:**
- Patients were recruited consecutively from a representative population with a control group recruited from the same population, methods and interventions were clearly defined to allow repetition, all important confounders were identified and adjusted for in the analyses, and all patients were followed up for at least 1 year

**Derived cohorts from RCTs:**
- The cohort was derived from a RCT with appropriate randomisation and allocation concealment methods were used; patients, carers, investigators and outcome assessors blinded; ITT results reported; minimum of 1-year follow-up in all patients.
**Moderate:** the study is subject to bias/limitations, but these are unlikely to significantly impact on the reliability of the results

**RCTs:** appropriate randomisation and allocation concealment methods used but some imbalance at baseline in non-essential confounders; outcome assessors blinded; ITT results reported; mean/median follow-up at least 1 year

**Cohort studies:** patients were not recruited consecutively but were from a representative population with a control group recruited from the same population or patients were recruited consecutively from a representative population with no control group, methods and interventions were clearly defined to allow repetition, confounders were identified and adjusted for in the analyses, and the mean/median follow-up was 1 year or longer

**Derived cohorts:** the cohort was derived from a RCT where appropriate randomisation and allocation concealment methods were used; patients and outcome assessors blinded; ITT reported; mean/median follow-up at least 1 year

**Post hoc analyses:** the analysis was based on all the patients from a RCT with appropriate randomisation and allocation concealment methods were used; patients, carers, investigators and outcome assessors blinded; ITT results reported; minimum of 1-year follow-up in all patients

**Low:** the study is subject to bias/limitations, and these are likely to significantly impact on the reliability of the results

**RCTs:** the RCT fails to meet one of the essential criteria (appropriate randomisation or allocation concealment methods; blinding of patients and outcome assessors) or extreme imbalance at baseline in essential confounders, and follow-up was not at least 1 year in all patients

**Cohort studies:** patients were not recruited consecutively, there was no control group, important confounders were not identified and/or adjusted for in the analyses, and the minimum follow-up was not at least 1 year in all patients

**Derived cohorts:** the cohort was derived from a RCT that failed to meet one of the essential criteria (appropriate randomisation or allocation concealment methods; blinding of patients and outcome assessors) and follow-up was not at least 1 year in all patients

**Post hoc analyses:** the analysis was based on a subset of the patients from a RCT, all the patients from a RCT that would be considered as moderate or low quality, or patients from a cohort study

**Unclear:** there is insufficient information to judge the quality of the study

**BM, bone turnover marker.**
Appendix 3  Excluded studies with rationale

Systematic review of the clinical effectiveness evidence

Identified for full-paper screening 444; nine were unobtainable,196–204 (one200 was identified from a bibliography with the same journal details as a screened paper,205 but a different title), 35 were reviews that underwent bibliographic screening,5,6,11,15,34,47,50,69–96 and 42 were included. Of the remaining 358 studies, the reasons for exclusion are given below (some papers were excluded for more than one reason):

1. protocol linked to excluded trial: 4 (#206 linked to #207; #208 linked to #209; #210 linked to #211; #212 linked to #213 and #214)
2. abstract linked to an included study: 2107–111,113–122,124–129
3. abstracts linked to an excluded study: 6215–220
4. abstract with insufficient data to extract (authors contacted and either confirmed no further data available or did not reply): 6221–226
5. duplicate full paper of an included study: 5 (#58 linked to #110; 74 linked to #478; 195 linked to #125; #130 linked to #106; #123 linked to #137)
6. no data for osteoporotic and/or treated patients separately from a more heterogeneous population: 16105,134,227–239
7. manufacturer’s trial for which results are currently not available from the manufacturer’s database: 1240
8. no osteoporotic patients: 8337,181,202,205,241–319
9. patients who were not receiving an osteoporosis medication:
10. patients receiving an osteoporosis treatment but not one of the ones being evaluated in the review:
    1710,251,266,289,378–390
11. a bone turnover marker of interest not included in the tests used in the study:
12. not prospective: 1401
13. non-RCTs that had fewer than 20 patients reaching the analysis stage of the study: 1945,412–429
14. no outcomes of interest reported or sufficient data to calculate any:
    8531,46,51,52,59,100,102,103,105,124,138,214,221,225,227,228,231,240,430–496

Systematic review of the cost-effectiveness evidence

1. Examined health benefits only.167
2. Review article mentioned cost-effectiveness in the abstract, but reported only the characteristics and clinical effectiveness of bone turnover markers in the results; bibliography was scanned for relevant cost-effectiveness studies but none was cited.497
## Appendix 4  Data extraction tables

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Armstrong (2007)</strong></td>
<td><strong>Original study design:</strong> uncontrolled cohort</td>
<td><strong>Test 1:</strong> sCTX</td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review:</strong> uncontrolled cohort</td>
<td><strong>Assay method used:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis:</strong> NR</td>
<td><strong>Timing of test:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied:</strong> none reported</td>
<td><strong>Dietary restrictions:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given:</strong> all except four patients</td>
<td><strong>Time of collection:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Treatment:</strong> alendronate or risedronate regimen NR</td>
<td><strong>Storage temperature:</strong> NR</td>
</tr>
<tr>
<td></td>
<td>( N = 46; n \text{ with OP} = 46; n \text{ PMW} = \text{unclear}; n \text{ male} = 6 )</td>
<td><strong>Delay to freezing:</strong> NR</td>
</tr>
<tr>
<td></td>
<td>( n \text{ with prior fracture} = 20 )</td>
<td><strong>Time in storage:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements:</strong> NR</td>
<td><strong>Specialist laboratory:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements:</strong> NR</td>
<td><strong>LSC:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up:</strong> NR</td>
<td><strong>Equation:</strong> NR</td>
</tr>
<tr>
<td><strong>Bauer (2004)</strong></td>
<td><strong>Original study design:</strong> RCT</td>
<td><strong>Intra-assay CV:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review:</strong> derived single-arm cohort(s)</td>
<td><strong>Inter-assay CV:</strong> 12%</td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis:</strong> T-score (location unspecified) ( \leq -2.5 ); vertebral fracture</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied:</strong> none reported</td>
<td><strong>Number of:</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given:</strong> only those with deficiency</td>
<td><strong>samples:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Treatment:</strong> alendronate 5–10 mg/day at second annual visit orally for 2 years</td>
<td><strong>replicates per run:</strong> NR</td>
</tr>
<tr>
<td></td>
<td>( N = 3105; n \text{ with OP} = 3105; n \text{ PMW} = 3105; n \text{ male} = 0 )</td>
<td><strong>Analytical sensitivity:</strong> NR</td>
</tr>
<tr>
<td></td>
<td>( n \text{ with prior fracture: vertebral: 1022; non-vertebral: 819} )</td>
<td><strong>Upper normal limit:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements:</strong> mean spine: g/cm(^2): 0.83; hip: g/cm(^2): 0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements:</strong> sCTX: 3327 pmol/l; sBALP: 13.7 ng/ml; sP1NP: 51.4 ng/ml</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up:</strong> mean: 3.6; range 2.5 to 4.5 years</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Population and treatment details</td>
<td>Intervention/test details</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td><strong>Test 3</strong>: sCTX</td>
<td><strong>Assay method used</strong>: ELISA</td>
<td><strong>Number of</strong>:</td>
</tr>
<tr>
<td></td>
<td><strong>Timing of test</strong>: baseline; each annual visit</td>
<td><strong>samples</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Dietary restrictions</strong>: at baseline 20% fasted – otherwise none</td>
<td><strong>replicates per run</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Storage temperature</strong>: –20 °C for approx. 3 years, then –70 °C</td>
<td><strong>Analytical sensitivity</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Delay to freezing</strong>: NR</td>
<td><strong>Upper normal limit</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Time in storage</strong>: &gt; 28 days (up to 8.7 years)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Specialist laboratory</strong>: yes</td>
<td><strong>Equation</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>LSC</strong>: NR</td>
<td><strong>Intra-assay CV</strong>: 5%</td>
</tr>
<tr>
<td></td>
<td><strong>Intra-assay CV</strong>: 8%</td>
<td><strong>Inter-assay CV</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Number of</strong>:</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>samples</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>replicates per run</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Analytical sensitivity</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Upper normal limit</strong>: NR</td>
<td></td>
</tr>
<tr>
<td><strong>Test 4</strong>: DXA</td>
<td><strong>Area assessed</strong>: hip (unspecified); posteroanterior LS; units used: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Timing of test</strong>: annually</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>LSC</strong>: NR</td>
<td><strong>Equation</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Precision error</strong>: NR</td>
<td><strong>Number of technicians</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis</strong>: T-score at LS/FN ≤ –2.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied</strong>: conditions known to influence bone metabolism; lifestyle known to influence bone metabolism; liver dysfunction; medications known to affect bone metabolism; renal impairment and/or transplant</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given</strong>: everyone</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Treatment</strong>: raloxifene 60 or 120 mg/day (duration NR)</td>
<td><strong>Follow-up</strong>: NR</td>
</tr>
<tr>
<td></td>
<td>N = NR; n with OP = NR; n PMW = NR; n male = 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Mean age</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n with prior fracture: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements</strong>: NR</td>
<td></td>
</tr>
<tr>
<td><strong>Test 1</strong>: sBALP</td>
<td><strong>Assay method used</strong>: IRMA</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Timing of test</strong>: baseline; 6, 12, 24 and 36 months</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Dietary restrictions</strong>: 6-hour fast</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Time of collection</strong>: any time of day after fasting</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Storage temperature</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Delay to freezing</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Time in storage</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Specialist laboratory</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>LSC</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Equation</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Precision error</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Number of technicians</strong>: NR</td>
<td></td>
</tr>
<tr>
<td><strong>Test 2</strong>: uCTX</td>
<td><strong>Assay method used</strong>: ELISA</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Timing of test</strong>: baseline; 6, 12, 18, 24 and 36 months</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Sample type</strong>: first morning void; Corrected for Cr: yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Dietary restrictions</strong>: none</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Storage temperature</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Delay to freezing</strong>: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Time in storage</strong>: NR</td>
<td></td>
</tr>
</tbody>
</table>
### Study Population and treatment details

<table>
<thead>
<tr>
<th>Test</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| **Test 3** | DXA  
Area assessed: FN; LS (L1–L4); units used: g/cm²  
Timing of test: baseline; 12 and 24 months |
| **Test 1** | sBALP  
Assay method used: chemiluminescence  
Timing of test: baseline; 6 months  
Dietary restrictions: NR  
Time of collection: between 07.00 and 16.00  
Storage temperature: −20 °C, then −80 °C  
Delay to freezing: NR  
Time in storage: up to 4 months prior to dispatch to laboratory and storage at −80 °C  
Specialist laboratory: NR  
LSC: NR  
Equation: NR  
Precision error: NR  
Number of technicians: NR |
| **Test 2** | sP1NP  
Assay method used: electrochemiluminescence  
Timing of test: baseline; 6 months  
Dietary restrictions: NR  
Time of collection: between 07.00 and 16.00  
Storage temperature: −20 °C, then −80 °C  
Delay to freezing: NR  
Time in storage: up to 4 months prior to dispatch to laboratory and storage at −80 °C  
Specialist laboratory: yes  
LSC: NR  
Equation: NR  
Precision error: NR  
Number of technicians: NR |

### Original study design: RCT

**Study design as used in this review:** derived single-arm cohort(s)  
**Definition of osteoporosis:** T-score at LS/hip ≤ −2.5 + ≥ 1 OP fracture  
**Exclusion criteria applied:** conditions known to influence bone metabolism; contraindications to treatment; hypersensitivity to bisphosphonate; medications known to affect bone metabolism  
**Supplemental Ca or vitamin D given:** everyone  
**Treatment:** teriparatide 20 µg/day SC for 1 or 2 years  
N = 758; n with OP = 758; n PMW = 758; n male = 0  
Mean age: 69.8 (SD 7.5) years  
n with prior fracture: NR  
Baseline BMD measurements: mean LS: g/cm²: 0.738 and T-score: −3.21; FN: g/cm²: 0.624  
Baseline BM measurements: mean  
Treatment-naïve group: BALP 12.9 µg/L; P1NP 48.2 µg/L  
Prior therapy group: BALP 10.1 µg/L; P1NP 26.1 µg/L  
Prior non-response to therapy group: BALP 10.2 µg/L; P1NP 27.5 µg/L  
**Follow-up:** range 12 to 24 months
**Study Population and treatment details**

**Number of:**
- samples: NR
- replicates per run: NR

**Analytical sensitivity:** NR
**Upper normal limit:** NR

**Test 3: DXA**
- **Area assessed:** FN; LS (L1–L4); Total hip; units used: g/cm²
- **Timing of test:** baseline; 6, 12, 18 and 24 months
- **LSC:** NR
- **Equation:** NR
- **Precision error:** NR

**Number of technicians:** quality assessment and evaluation by a central reader

---

**Bruyere (2010),¹⁶¹ western Europe (original RCTs multinational) English Study dates: NR Full published paper**

**Original study design:** post hoc subgroup analysis of patients from two RCTs combined

**Study design as used in this review:** derived single-arm cohort(s)

**Definition of osteoporosis:** T-score (location unspecified) ≤−2.5 + ≥ 1 risk factor for fracture

**Exclusion criteria applied:** conditions known to influence bone metabolism; medications known to affect bone metabolism

**Supplemental Ca or vitamin D given:** only those with deficiency

**Treatment:** Strontium ranelate 2 g/day orally (duration NR)
- N = 2373; n with OP = 2373; n PMW = 2373; n male = 0
- Mean age: 73.9 (SD 6.1) years
- n with prior fracture: NR

**Baseline BMD measurements:** mean LS T-score: −3.06; FN T-score: −2.99

**Baseline BM measurements:** median sCTX: 0.509 ng/ml; sBALP: 11.3 ng/ml; uNTX: 49.6 nM BCE/mM Cr

**Follow-up:** maximum 3 years

---

**Test 1: sBALP**
- **Assay method used:** IRMA
- **Timing of test:** baseline; 3 months
- **Dietary restrictions:** fasting (details NR)
- **Time of collection:** NR
- **Storage temperature:** −80 °C
- **Delay to freezing:** NR
- **Specialist laboratory:** yes
- **LSC:** NR
- **Equation:** NR
- **Intra-assay CV:** maximum: 10%
- **Inter-assay CV:** maximum: 10%
- **Number of:**
  - samples: NR
  - replicates per run: NR
- **Analytical sensitivity:** 2 ng/ml
- **Upper normal limit:** NR

**Test 2: sCTX**
- **Assay method used:** ELISA
- **Timing of test:** baseline; 3 months
- **Dietary restrictions:** fasting (details NR)
- **Time of collection:** NR
- **Storage temperature:** −80 °C
- **Delay to freezing:** NR
- **Specialist laboratory:** yes
- **LSC:** NR
- **Equation:** NR
- **Intra-assay CV:** maximum: 10%
- **Inter-assay CV:** maximum: 10%
- **Number of:**
  - samples: NR
  - replicates per run: NR
- **Analytical sensitivity:** 0.016 ng/ml
- **Upper normal limit:** NR

**Test 3: uNTX**
- **Assay method used:** ELISA
- **Timing of test:** baseline; 3 months
- **Sample type:** second morning void Corrected for Cr: yes
- **Dietary restrictions:** NR
- **Storage temperature:** −80 °C
- **Delay to freezing:** NR
### Study Population and Treatment Details

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and Treatment Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original study design:</strong> post hoc subgroup analysis of a RCT</td>
<td></td>
</tr>
<tr>
<td><strong>Study design as used in this review:</strong> derived single-arm cohort(s)</td>
<td></td>
</tr>
<tr>
<td><strong>Definition of osteoporosis:</strong> T-score at LS/hip $\leq -1.0 + \geq 1$ OP fracture; T-score at LS/hip $\leq -2.0$</td>
<td></td>
</tr>
<tr>
<td><strong>Exclusion criteria applied:</strong> none reported</td>
<td></td>
</tr>
<tr>
<td><strong>Supplemental Ca or vitamin D given:</strong> everyone</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment 1:</strong> alendronate 10 mg/day orally for at least 18 months</td>
<td></td>
</tr>
<tr>
<td>$N = 77$; $n$ with OP $= 77$; $n$ PMW $= 50$; $n$ male $= 17$</td>
<td></td>
</tr>
<tr>
<td>Mean age: 60.6 (SE 2.5) years</td>
<td></td>
</tr>
<tr>
<td>$n$ with prior fracture: vertebral: 17; non-vertebral: 34</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline BMD measurements:</strong> mean LS T-score: $-2.7$; FN T-score: $-2.2$</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline BM measurements:</strong> sCTX: 3264 pmol/l; sBAP: 8.7 µg/l; sP1NP: 40 µg/l</td>
<td></td>
</tr>
<tr>
<td><strong>Follow-up:</strong> maximum 18 months</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment 2:</strong> Teriparatide 20 µg/day IM/SC – unclear for at least 18 months</td>
<td></td>
</tr>
<tr>
<td>$N = 80$; $n$ with OP $= 80$; $n$ PMW $= 41$; $n$ male $= 13$</td>
<td></td>
</tr>
<tr>
<td>Mean age: 56.1 (SE 2.6) years</td>
<td></td>
</tr>
<tr>
<td>$n$ with prior fracture: vertebral: 18; non-vertebral: 28</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline BMD measurements:</strong> mean LS T-score: $-2.5$; FN T-score: $-2.0$</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline BM measurements:</strong> sCTX: 3503 pmol/l; sBAP: 9.8 µg/l; sP1NP: 43 µg/l</td>
<td></td>
</tr>
<tr>
<td><strong>Follow-up:</strong> maximum 18 months</td>
<td></td>
</tr>
</tbody>
</table>

---

### Test 1: sBALP

- **Assay method used:** IRMA
- **Timing of test:** baseline; 1, 6 and 18 months
- **Dietary restrictions:** overnight/ morning fasting
- **Time of collection:** morning
- **Storage temperature:** NR
- **Delay to freezing:** NR
- **Number of:** samples: NR; replicates per run: NR
- **Analytical sensitivity:** NR
- **Upper normal limit:** NR

### Test 2: sP1NP

- **Assay method used:** RIA
- **Timing of test:** baseline; 1, 6 and 18 months
- **Dietary restrictions:** overnight/ morning fasting
- **Time of collection:** morning
- **Storage temperature:** NR
- **Delay to freezing:** NR
- **Time in storage:** NR
- **Specialist laboratory:** yes
- **LSC:** NR
- **Equation:** NR
- **Intra-assay CV:** range 3.2% to 5.2%
- **Inter-assay CV:** range 7.4% to 7.9%
- **Number of:** samples: NR; replicates per run: NR
- **Analytical sensitivity:** NR
- **Upper normal limit:** NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chen (2005)</strong></td>
<td><strong>Test 3</strong>: sCTX</td>
<td><strong>Dietary restrictions</strong>: overnight/morning fasting; <strong>Timing of test</strong>: baseline; 1, 6 and 18 months; <strong>Assay method used</strong>: ELISA; <strong>Storage temperature</strong>: NR; <strong>Delay to freezing</strong>: NR; <strong>Time in storage</strong>: NR; <strong>Specialist laboratory</strong>: yes; <strong>Equation</strong>: NR; <strong>Inter-assay CV</strong>: range 11.1% to 13.5%</td>
</tr>
<tr>
<td></td>
<td><strong>Test 1</strong>: sBALP</td>
<td><strong>Timing of test</strong>: baseline; 1, 3, 6 and 12 months; study end; <strong>Assay method used</strong>: IRMA; <strong>Storage temperature</strong>: –20 °C; <strong>Delay to freezing</strong>: NR; <strong>Time in storage</strong>: 4 months; <strong>Specialist laboratory</strong>: yes; <strong>Equation</strong>: NR; <strong>Inter-assay CV</strong>: range 7.4% to 7.9%</td>
</tr>
<tr>
<td></td>
<td><strong>Test 2</strong>: sP1NP</td>
<td><strong>Timing of test</strong>: baseline; 6, 12 and 18 months; <strong>LSC</strong>: NR; <strong>Equation</strong>: NR; <strong>Precision error</strong>: NR; <strong>Number of technicians</strong>: NR</td>
</tr>
<tr>
<td>Study</td>
<td>Population and treatment details</td>
<td>Intervention/test details</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td></td>
<td>FN 0.64 units NR</td>
<td><strong>Baseline BM</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Assay method used</strong>: RIA</td>
<td><strong>Timing of test</strong>: baseline; 3 months</td>
</tr>
<tr>
<td></td>
<td><strong>Dietary restrictions</strong>: NR</td>
<td><strong>Time of collection</strong>: morning</td>
</tr>
<tr>
<td></td>
<td><strong>Storage temperature</strong>: –20 °C</td>
<td><strong>Delay to freezing</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Time in storage</strong>: NR</td>
<td><strong>Specialist laboratory?</strong>: Yes</td>
</tr>
<tr>
<td></td>
<td><strong>LSC</strong>: NR</td>
<td><strong>Equation</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Intra-assay CV</strong>: NR</td>
<td><strong>Inter-assay CV</strong>: range 3.1% to 8.2%</td>
</tr>
<tr>
<td></td>
<td><strong>Number of</strong>:</td>
<td><strong>samples</strong>: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>replicates per run</strong>: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Analytical sensitivity</strong>: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Upper normal limit</strong>: NR</td>
</tr>
</tbody>
</table>

**Test 3: uNTX**

- **N** = 520; **n with OP** = 520;
- **n male** = 0; **n with prior fracture**: NR
- **Mean age**: 69 (SD 6.9) years
- **Baseline BMD**: mean LS 0.82;
- **Baseline BM**: mean FN 0.64 units NR

**Baseline BM**: NR

- **Assay method used**: ELISA
- **Timing of test**: baseline; 1, 3, 6 and 12 months; study end
- **Sample type**: NR
- **Corrected for Cr**: yes
- **Dietary restrictions**: NR
- **Time of collection**: morning
- **Storage temperature**: –20 °C
- **Delay to freezing**: NR
- **Time in storage**: 1 year
- **Specialist laboratory**: Yes
- **LSC**: NR
- **Equation**: NR
- **Intra-assay CV**: NR
- **Inter-assay CV**: range 6.7% to 14.8%
- **Number of**: samples: NR, replicates per run: NR
- **Analytical sensitivity**: NR
- **Upper normal limit**: NR

**Test 4: DXA**

- **Area assessed**: FN; LS (unspecified);
- **units used**: g/cm²
- **Timing of test**: baseline; 12 and 18 months
- **LSC**: 3%
- **Equation**: NR
- **Precision error**: –3% to 3%
- **Number of technicians**: NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| Clowes (2003), UK, 147 | **Study design as used in this review:** derived single-arm cohort(s) | **Test 1:** sP1NP  
**Assay method used:** NR  
**Timing of test:** baseline; 1, 2, 4, 8, 12, 24, 25 weeks  
**Dietary restrictions:** overnight fasting  
**Time of collection:** morning  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** 1.6  
**Number of samples:** NR  
**Number of replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR |
|               | **Definition of osteoporosis:** NR |                           |
|               | **Exclusion criteria applied:** none reported |                           |
|               | **Supplemental Ca or vitamin D given:** Ca – everyone |                           |
|               | **Treatment:** Raloxifene 60 mg/day orally (duration NR)  
N = 22; n with OP = 22; n PMW = unclear; n male = unclear  
Mean age: NR  
N with prior fracture: NR  
Baseline BMD measurements: NR  
Baseline BM measurements: NR |                           |
|               | **Follow-up:** maximum 25 weeks |                           |
| Delmas (2007), multinational, 56 | **Original study design:** RCT (cluster) | **Test 2:** sCTX  
**Assay method used:** electrochemiluminescence  
**Timing of test:** baseline; 1, 2, 4, 8, 12, 24, 25 weeks  
**Dietary restrictions:** overnight fasting  
**Time of collection:** morning  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** 4.7  
**Number of samples:** NR  
**Number of replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR |
|               | **Study design as used in this review:** RCT |                           |
|               | **Definition of osteoporosis:** T-score at LS/hip ≤ –2.5 |                           |
|               | **Exclusion criteria applied:** medications known to affect bone metabolism |                           |
|               | **Supplemental Ca or vitamin D given:** everyone |                           |
|               | **Treatment:** risendronate 5 mg/day orally for 1 year  
N = 2382; n with OP = 2382; n PMW = 2382; n male = 0  
Mean age: NR  
N with prior fracture: NR  
Baseline BMD measurements: NR  
Baseline BM measurements: NR |                           |
|               | **Follow-up:** maximum 12 months |                           |
|               | **Intervention:** BM feedback  
(uNTX at 13 and 25 weeks)  
N = 1189; n with OP = 1189; n male = 0; n with prior vertebral fracture: 359  
Mean age: 71.1 (SD 4.3) years  
Baseline BMD: mean spine T-score –2.8; hip T-score –2.0  
Baseline BM: NR |                           |
|               | **Intervention:** no BM feedback  
N = 1113; n with OP = 1113; n male = 0; n with prior vertebral fracture: 330  
Mean age: 71.5 (SD 4.5) years  
Baseline BMD: mean spine T-score –2.8; hip T-score –2.0  
Baseline BM: NR |                           |
|               | **Test:** uNTX  
**Assay method used:** ELISA  
**Timing of test:** baseline; 10 and 22 weeks  
**Sample type:** second morning void  
**Corrected for Cr:** NR |                           |
**Study**

Delmas (2009), multinational

**English**

Study dates: NR

Full published paper

**Original study design**: post hoc subgroup analysis of an RCT

**Study design as used in this review**: derived single-arm cohort(s)

**Definition of osteoporosis**: T-score ≤ –1.5 + 1 moderate or 2 mild vertebral fractures; T-score at LS/hip ≤ –2.5

**Exclusion criteria applied**: conditions known to influence bone metabolism; medications known to affect bone metabolism; renal impairment and/or transplant; previous bisphosphonate use not in accordance with the washout schedule

**Supplemental Ca or vitamin D given**: everyone

**Treatment**: zoledronate 5 mg/year i.v. for 3 years

*N = NR; n male = 0*

**Mean age**: NR

*n with prior fracture**: NR

**Baseline BMD measurements**: NR

**Baseline BM measurements**: NR

**Follow-up**: maximum 36 months

<table>
<thead>
<tr>
<th>Test 1: sBALP</th>
<th><strong>Intervention/test details</strong></th>
</tr>
</thead>
</table>
| *N = 299; n with OP = 299; n male = 0; n with prior vertebral fracture = 173* | **Dietary restrictions**: NR  
**Storage temperature**: NR  
**Delay to freezing**: NR  
**Time in storage**: NR  
**Specialist laboratory**: yes  
**LSC**: 30  
**Equation**: NR  
**Intra-assay CV**: range: 1.1% to 6.7%  
**Inter-assay CV**: range 3.5% to 7.8%  
**Number of samples**: NR  
**Replicates per run**: NR  
**Analytical sensitivity**: NR  
**Upper normal limit**: NR |

**Timing of test**: baseline; 6, 12 and 18 months; and 1, 3, 6 and 12 months after the third (final) infusion

**Dietary restrictions**: fasting (details NR)

**Time of collection**: NR

**Storage temperature**: NR

**Delay to freezing**: NR

**Time in storage**: NR

**Specialist laboratory**: Yes

**LSC**: NR

**Equation**: NR

**Intra-assay CV**: range: 2.3 to 3.7

**Inter-assay CV**: range: 4.4 to 9.8

**Number of samples**: NR

**Replicates per run**: NR

**Analytical sensitivity**: NR

**Upper normal limit**: NR

<table>
<thead>
<tr>
<th>Test 2: sP1NP</th>
<th><strong>Intervention/test details</strong></th>
</tr>
</thead>
</table>
| *N = 618; n with OP = 618; n with prior vertebral fracture = 361* | **Dietary restrictions**: fasting (details NR)  
**Time of collection**: NR  
**Storage temperature**: NR  
**Delay to freezing**: NR  
**Time in storage**: NR  
**Specialist laboratory**: Yes  
**LSC**: NR  
**Equation**: NR  
**Intra-assay CV**: range: 2.3 to 3.7

**Timing of test**: baseline; 6, 12 and 18 months; and 1, 3, 6 and 12 months after the third (final) infusion

**Dietary restrictions**: fasting (details NR)

**Time of collection**: NR

**Storage temperature**: NR

**Delay to freezing**: NR

**Time in storage**: NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Test 3</strong>: sCTX</td>
</tr>
<tr>
<td>Test 3</td>
<td>N = 299; n with OP = 299; n male = 0; n with prior vertebral fracture = 173</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 74.8 (SD 5.8) years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline BMD: mean FN 0.54 -2.785</td>
<td>Baseline BM: 0.36 ng/ml</td>
</tr>
<tr>
<td></td>
<td>Assay method used: electrochemiluminescence</td>
<td>Assay method used:</td>
</tr>
<tr>
<td></td>
<td>Timing of test: baseline; 6, 12 and 18 months; 1, 3, 6 and 12 months after the third (final) infusion</td>
<td>electrochemiluminescence</td>
</tr>
<tr>
<td></td>
<td>Dietary restrictions: fasting</td>
<td>Timing of test:</td>
</tr>
<tr>
<td></td>
<td>(details NR)</td>
<td>baseline; 6, 12, 24 and 36 months</td>
</tr>
<tr>
<td></td>
<td>Intra-assay CV: range: 1.1% to 6.7%</td>
<td>Diet restrictions:</td>
</tr>
<tr>
<td></td>
<td>Inter-assay CV: range 3.8% to 6.1%</td>
<td>fasting (details NR)</td>
</tr>
<tr>
<td></td>
<td>Number of:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>samples: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>replicates per run: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analytical sensitivity: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper normal limit: NR</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test 4</th>
<th>DXA</th>
<th>Area assessed: FN; Units used: g/cm²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Timing of test: baseline; 6, 12, 24 and 36 months</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LSC: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Equation: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Precision error: NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of technicians: 1</td>
</tr>
<tr>
<td>Study</td>
<td>Population and treatment details</td>
<td>Intervention/test details</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
| Dobnig (2005),¹⁵ multinational English Study dates: NR Full published paper | **Original study design:** post hoc subgroup analysis of an RCT  
**Study design as used in this review:** derived single-arm cohort(s)  
**Definition of osteoporosis:** one moderate or two mild vertebral fractures; T-score at LS/hip \(\leq -1.0 + \geq 1\) OP fracture  
**Exclusion criteria applied:** patients without a biopsy with at least one 2D or 3D microCT from the specimen  
**Supplemental Ca or vitamin D given:** everyone  
**Treatment:** Teriparatide 20 or 40 µ/day IM/SC (NR which used) for a median of 20 (range 17 to 22 months)  
\(N = 36; n\ with\ OP = 36; n\ PMW = 36; n\ male = 0\)  
**Mean age:** 67.9 (SD 6.2) years  
\(n\ with\ prior\ fracture: NR\)  
**Baseline BMD measurements:** mean LS: g/cm²: 0.8; FN g/cm²: 0.63  
**Baseline BM measurements:** mean sBALP: 14.7 µg/l; uNTX: 45.6 nM BCE/mM  
**Follow-up:** Mean: 22 months (range 19 to 25)  
**Test 1:** sBALP  
**Assay method used:** IRMA  
**Timing of test:** baseline; 1, 3, 6 and 12 months; study end  
**Dietary restrictions:** overnight fasting, Time of collection: morning  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** range: 4.2% to 6.8%  
**Inter-assay CV:** range 7.4% to 7.9%  
**Number of:**  
   samples: NR  
   replicates per run: NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR  
**Test 2:** uNTX  
**Assay method used:** ELISA  
**Timing of test:** baseline; 1, 3, 6 and 12 months; study end  
**Sample type:** second morning void; Corrected for CR: yes  
**Dietary restrictions:** overnight fasting  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** range: 4.5% to 6.6%  
**Inter-assay CV:** range 6.7% to 14.8%  
**Number of:**  
   samples: NR  
   replicates per run: NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR  
**Test 3:** Biopsy  
**Site:** iliac crest  
**Number:** NR  
**Needle:** NR  
**Technique:** bordier  
**Embedding method:** NR  
**Anaesthesia:** NR  
**Number of clinicians:** NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| **Dobnig** (2006), 153 western Europe English Study dates: NR Full published paper | **Original study design:** RCT | **Test 1:** sCTX  
Assay method used: ELISA  
Timing of test: baseline; 2, 6 and 12 months  
Dietary restrictions: overnight fasting  
Time of collection: morning (08.00 to 10.00)  
Storage temperature: NR  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: NR  
LSC: NR  
Equation: NR  
Intra-assay CV: 9%  
Inter-assay CV: 10%  
Number of: samples: NR  
replicates per run: NR  
Analytical sensitivity: LLOD 1.2 pg/ml; Upper normal limit: NR |
| | **Study design as used in this review:** derived single-arm cohort(s) | **Test 2:** DXA  
Area assessed: FN; Units used: Z- and T-scores  
Timing of test: baseline; 12 months  
LSC: NR  
Equation: NR  
Precision error: 1.6%  
Number of technicians: NR |
| | **Definition of osteoporosis:** T-score at LS/hip ≤ –2.5  
Exclusion criteria applied: < 60 years old; conditions known to influence bone metabolism; medications known to affect bone metabolism  
Supplemental Ca or vitamin D given: everyone | **Treatment:** alendronate 10 mg/day or risedronate 5 mg/day orally (duration NR)  
N = 37; n with OP = 37; n PMW = 37; n male = 0  
Mean age: 69 (SD 4.0) years  
N with prior fracture: NR  
Baseline BMD measurements: FN Z-score –1.03  
Baseline BM measurements: sCTX: 2.58 nmol/l  
Follow-up: maximum 12 months |
| **Eastell** (2003), 137 multinational English Study dates: NR Full published paper | **Original study design:** post hoc subgroup analysis of two RCTs combined | **Test 1:** uCTX  
Assay method used: ELISA  
Timing of test: baseline; 3 and 6 months  
Sample type: second morning void  
Corrected for Cr: yes  
Dietary restrictions: NR  
Storage temperature: –20 °C  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: NR  
LSC: NR  
Equation: NR  
Intra-assay CV: 9%  
Inter-assay CV: 10%  
Number of: samples: NR  
replicates per run: NR  
Analytical sensitivity: LLOD 1.2 pg/ml; Upper normal limit: NR |
| | **Study design as used in this review:** derived single-arm cohort(s) | **Test 2:** uNTX  
Assay method used: chemiluminescence  
Timing of test: baseline; 3 and 6 months  
Sample type: second morning void  
Corrected for Cr: yes  
Dietary restrictions: NR  
Storage temperature: –20 °C; |
| | **Definition of osteoporosis:** at least two vertebral fractures; T-score (location unspecified) ≤ –2.0 + ≥ 2 OP fracture  
Exclusion criteria applied: none reported  
Supplemental Ca or vitamin D given: everyone | **Treatment:** risedronate 5 mg/day orally for 3 years  
N = 358; n with OP = 358; n PMW = 358; n male = 0  
Mean age: 70 (SD 7.8) years  
N with prior fracture: vertebral: 324  
Baseline BMD measurements: mean LS T-score: 2.53; FN T-score: 2.27  
Baseline BM measurements: median uCTX: 7.36 nmol/nmol; uNTX: 68.6 nmol BCE/mmol  
Follow-up: maximum 3 years |
### Study Population and treatment details

<table>
<thead>
<tr>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay to freezing: NR</td>
</tr>
<tr>
<td>Time in storage: NR</td>
</tr>
<tr>
<td>Specialist laboratory: NR</td>
</tr>
<tr>
<td>LSC: NR</td>
</tr>
<tr>
<td>Equation: NR</td>
</tr>
<tr>
<td>Intra-assay CV: NR</td>
</tr>
<tr>
<td>Inter-assay CV: maximum 6.7%</td>
</tr>
<tr>
<td>Number of:</td>
</tr>
<tr>
<td>samples: NR</td>
</tr>
<tr>
<td>replicates per run: NR</td>
</tr>
<tr>
<td>Analytical sensitivity: NR</td>
</tr>
<tr>
<td>Upper normal limit: NR</td>
</tr>
</tbody>
</table>

### Test 3: DXA

- **Area assessed:** FN; LS (L1–L4);
- **T-score**
- **Timing of test:** baseline; 12 and 36 months
- **LSC:** NR
- **Equation:** NR
- **Precision error:** NR
- **Number of technicians:** NR

### Original study design

- **Post hoc subgroup analysis of a RCT**

### Study design as used in this review

- **Derived single-arm cohort(s)**

### Definition of osteoporosis

- **T-score at LS/hip ≤ −2.5**

### Exclusion criteria applied

- < 60 years old; conditions known to influence bone metabolism; medications known to affect bone metabolism; serum vitamin D level < 12 ng/ml; T-score < −4.0; prior BPs (unless < 3 years and 12 months without treatment prior to entry into study), IV BPs, PTH or its derivatives
- **Supplemental Ca or vitamin D given:** everyone

### Treatment

- **Denosumab 60 mg every 6 months SC for 3 years**
- **N = 96; n with OP = 96; n PMW = 96; n male = 0**
- **Mean age:** 72.3 (SD 5.0) years
- **n with prior fracture:** vertebral: 23
- **Baseline BMD measurements:** mean LS T-score: −2.88; hip T-score: −1.93

### Follow-up

- **Maximum 36 months**

---

**References**

- Eastell (2011). From the original study.
- **Multinational English**
- **Study dates:** NR
- **Full published paper:**

---

**Notes:**

- **DOI:** 10.3310/hta18110
- **Health Technology Assessment 2014 Vol. 18 No. 11**
- **© Queen’s Printer and Controller of HMSO 2014.** This work was produced by Burch et al. under the terms of a commissioning contract issued by the Secretary of State for Health.
- **This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.**
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test 3</strong>: sCTX&lt;br&gt;<strong>Assay method used</strong>: ELISA&lt;br&gt;<strong>Timing of test</strong>: baseline; 1, 6, 12, 24 and 36 months&lt;br&gt;<strong>Dietary restrictions</strong>: fasting (details NR)&lt;br&gt;<strong>Time of collection</strong>: NR&lt;br&gt;<strong>Storage temperature</strong>: NR&lt;br&gt;<strong>Delay to freezing</strong>: NR&lt;br&gt;<strong>Time in storage</strong>: NR&lt;br&gt;<strong>Specialist laboratory</strong>: NR&lt;br&gt;<strong>Reference interval</strong>: 0.2 to 0.9 ng/ml&lt;br&gt;<strong>LSC</strong>: NR&lt;br&gt;<strong>Equation</strong>: NR&lt;br&gt;<strong>Intra-assay CV</strong>: NR&lt;br&gt;<strong>Inter-assay CV</strong>: NR&lt;br&gt;<strong>Analytical sensitivity</strong>: LLOQ 0.049 ng/ml&lt;br&gt;<strong>Upper normal limit</strong>: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test 1</strong>: sP1NP (intact)&lt;br&gt;<strong>Assay method used</strong>: RIA – manual&lt;br&gt;<strong>Timing of test</strong>: baseline; 3 months&lt;br&gt;<strong>Dietary restrictions</strong>: fasting (details NR)&lt;br&gt;<strong>Time of collection</strong>: morning (07.30–10.00)&lt;br&gt;<strong>Storage temperature</strong>: −70 °C&lt;br&gt;<strong>Delay to freezing</strong>: NR&lt;br&gt;<strong>Time in storage</strong>: NR&lt;br&gt;<strong>Specialist laboratory</strong>: yes&lt;br&gt;<strong>LSC</strong>: −20%&lt;br&gt;<strong>Equation</strong>: 1.96 × sqrt(2 × CVi);&lt;br&gt;<strong>Intra-assay CV</strong>: range: 3.7% to 5.0%;&lt;br&gt;<strong>Number of</strong>: samples: 3; replicates per run: 20&lt;br&gt;<strong>Inter-assay CV</strong>: range 4.1% to 7.6%;&lt;br&gt;<strong>Number of</strong>: samples: 4; replicates per run: 20 different runs&lt;br&gt;<strong>Analytical sensitivity</strong>: 1 µg/l&lt;br&gt;<strong>Upper normal limit</strong>: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test 4</strong>: DXA&lt;br&gt;<strong>Area assessed</strong>: hip (unspecified); LS (unspecified); units used: NR&lt;br&gt;<strong>Timing of test</strong>: baseline; 12, 24 and 36 months;&lt;br&gt;<strong>LSC</strong>: NR&lt;br&gt;<strong>Equation</strong>: NR&lt;br&gt;<strong>Precision error</strong>: NR&lt;br&gt;<strong>Number of technicians</strong>: NR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Garnero (2008)**<br>**USA/Canada**<br>**English**<br>**Study design as used in this review**: derived single-arm cohort(s)

**Definition of osteoporosis**: T-score at LS/hip ≤ −2.0 + 1 risk factor; T-score at LS/hip ≤ −2.5

**Exclusion criteria applied**: conditions known to influence bone metabolism; medications known to affect bone metabolism; women who had been treated with bisphosphonates for > 12 months or for > 4 weeks during the previous 12 months

**Supplemental Ca or vitamin D given**: everyone

**Treatment**: alendronate 10 mg/day orally for 3 months<br>N = 60; n with OP = 60; n PMW = 60; n male = 0<br>Mean age: 70.7 (SD 6.8) years<br>n with prior fracture: 25<br>Baseline BMD measurements: Mean LS: g/cm²: 0.778; FN: g/cm²: 0.596<br>Baseline BM measurements: NR

**Follow-up**: minimum 12 months
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| Heaney (2011), 162 USA/Canada | **Original study design**: post hoc subgroup analysis of a RCT (different forms of calcium – everyone got PTH)  
**Study design as used in this review**: derived single-arm cohort(s)  
**Definition of osteoporosis**: T-score at LS/hip ≤ -1.0 + ≥ 1 OP fracture  
**Exclusion criteria applied**: < 60 years old; patients not adherent to teripertide treatment  
**Supplemental Ca or vitamin D given**: Ca – everyone (either as carbonate or as triphosphate)  
**Treatment**: teriparatide 20 µg/day SC (duration NR)  
\( N = 203; n \) with OP = 203; \( n \) PMV = 203; \( n \) male = 0  
Mean age: 70 (SD 6.67) years  
\( n \) with prior fracture: vertebral: 203  
Baseline BMD measurements: mean LS g/cm²: 0.866; hip: g/cm²: 0.722  
Baseline BM measurements: Mean uNTX: 32 nM BCE/mM | **Test 3**: sCTX  
**Assay method used**: electrochemiluminescence  
**Timing of test**: baseline; 3 months  
**Dietary restrictions**: NR  
**Time of collection**: NR  
**Storage temperature**: NR  
**Delay to freezing**: NR  
**Time in storage**: NR  
**Specialist laboratory**: NR  
**LSC**: 26% to 53%  
Equation: \( 1.96 \times \sqrt{2 \times \text{intraindividual CV}} \)  
**Intra-assay CV**: maximum: 4.1%  
**Inter-assay CV**: maximum: 5.7%  
**Number of**:  
\( \text{samples} \): NR  
\( \text{replicates per run} \): NR  
**Analytical sensitivity**: 0.01 ug/l;  
**Upper normal limit**: NR |
| **Follow-up**: maximum 12 months | **Test 1**: uNTX  
**Assay method used**: chemiluminescence  
**Timing of test**: baseline; 3, 6 and 12 months  
**Sample type**: 2 hour  
Corrected for Cr: yes  
**Dietary restrictions**: overnight/ morning fasting  
**Time of collection**: morning  
**Specialist laboratory**: NR  
**LSC**: NR  
Equation: NR  
**Intra-assay CV**: NR  
**Inter-assay CV**: NR  
**Number of**:  
\( \text{samples} \): NR  
\( \text{replicates per run} \): NR  
**Analytical sensitivity**: NR  
**Upper normal limit**: NR |
| **Test 2**: DXA  
**Area assessed**: hip (unspeciﬁed); LS (unspeciﬁed); units used: g/cm²  
**Timing of test**: baseline; 3, 6 and 12 months  
**LSC**: NR  
Equation: NR  
**Precision error**: NR  
**Number of technicians**: NR |
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| **Majima (2008)**<sup>160</sup> | **Original study design:** uncontrolled cohort | **Test 1:** sBAP  
**Assay method used:** ELISA  
**Timing of test:** baseline; 3, 6 and 12 months  
**Dietary restrictions:** overnight fasting  
**Time of collection:** morning (before 09:00)  
**Storage temperature:** –20 °C  
**Delay to freezing:** none  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** NR  
**Number of samples:** NR  
**Replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR |
| Asia  
**Study dates:** 2004 to 2007  
**Full published paper** | **Study design as used in this review:** uncontrolled cohort  
**Definition of osteoporosis:** T-score at LS/hip ≤ –2.5  
**Exclusion criteria applied:** conditions known to influence bone metabolism; DVT; ectopic calcifications; lifestyle known to influence bone metabolism; lumbar fracture; medications known to affect bone metabolism; patients who cannot walk well unaided; renal impairment and/or transplant  
**Supplemental Ca or vitamin D given:** no  
**Treatment:** raloxifene 60 mg/day orally for 12 months  
**N = 63; n with OP = 63; n PMW = 63; n male = 0  
**Mean age:** 70.49 (SD 9.1) years  
**n with prior fracture:** NR  
**Baseline BMD measurements:** mean LS: g/cm²: 0.671 and T-score: –3.176; FN: g/cm²: 0.547 and T-score: –2.201  
**Baseline BM measurements:** sBAP: 32.9 U/L; sNTX: 19.52 nmol BCE/L  
**Follow-up:** minimum 12 months | **Test 2:** sNTX  
**Assay method used:** ELISA  
**Timing of test:** baseline; 3, 6 and 12 months  
**Dietary restrictions:** overnight fasting  
**Time of collection:** morning (before 9:00)  
**Storage temperature:** –20 °C  
**Delay to freezing:** none  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** NR  
**Number of samples:** NR  
**Replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR  
**Test conducted:** DXA  
**Area assessed:** distal 1/3 radius; FN; LS (Unspecified); trochanter; ulradistal radius; Ward’s triangle; units used: g/cm²  
**Timing of test:** baseline; 6 and 12 months  
**LSC:** NR  
**Equation:** NR  
**Precision error:** 0.43%  
**Number of technicians:** NR |
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| Hochberg (2010), 163 USA/Canada English Study dates: NR Full published paper | **Original study design:** post hoc subgroup analysis of a RCT  
**Study design as used in this review:** derived single-arm cohort(s)  
**Definition of osteoporosis:** T-score at LS/hip ≤ −2.5  
**Exclusion criteria applied:** < 1 dose of trial medication; BM not measured; protocol violations; T-score < −5.0  
**Supplemental Ca or vitamin D given:** NR  
**Treatment:** ibandronate 150 mg monthly orally for 1 year  
\[ N = 323; n \text{ with OP} = 323; n \text{ PMW} = 323; n \text{ male} = 0 \]  
**Mean age:** 65.8 (SD 6.6) years  
**n with prior fracture:** 149  
**Baseline BMD measurements:** mean LS: g/cm²: 0.75  
**Baseline BM measurements:** mean sCTX: 0.53 ng/ml; sBALP: 12.28 ng/ml  
**Follow-up:** minimum 12 months | **Test 1:** sCTX  
**Assay method used:** electrochemiluminescence  
**Timing of test:** baseline; 3, 6 and 12 months  
**Dietary restrictions:** overnight fasting  
**Time of collection:** between 08:00 and 10:00  
**Storage temperature:** −80 °C  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** yes  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** NR  
**Number of:**  
  **samples:** NR  
  **replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR |
| Imai (2009), 136 Asia English Study dates: NR Full published paper | **Original study design:** uncontrolled cohort  
**Study design as used in this review:** uncontrolled cohort  
**Definition of osteoporosis:** LS BMD ≤ 70% young adult mean; vertebral fracture  
**Exclusion criteria applied:** < 49 years old; conditions known to influence bone metabolism; lumbar fracture; medications known to affect bone metabolism  
**Supplemental Ca or vitamin D given:** NR  
**Treatment:** alendronate 5 mg/day orally for 1 year  
\[ N = 37; n \text{ with OP} = 37; n \text{ PMW} = 37; n \text{ male} = 0 \]  
**Mean age:** 76.5 (SD 5.4) years  
**n with prior fracture:** NR  
**Baseline BMD measurements:** NR  
**Baseline BM measurements:** NR  
**Follow-up:** minimum 12 months | **Test 1:** uNTX  
**Assay method used:** NR  
**Sample type:** NR  
**Corrected for Cr:** NR  
**Timing of test:** baseline; 3 months  
**Dietary restrictions:** NR  
**Time of collection:** NR  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** NR  
**Number of:**  
  **samples:** NR  
  **replicates per run:** NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR | **Test 2:** DXA  
**Area assessed:** FN; LS (unspecified); total hip; units used: g/cm²  
**Timing of test:** baseline; 12 months  
**LSC:** 3% (also used 0% as BMD response in analyses);  
**Equation:** NR  
**Precision error:** 1%  
**Number of technicians:** NR |

**Test 2:** DXA  
**Area assessed:** LS (L2–L4); total hip; units used: g/cm²  
**Timing of test:** baseline; 6 and 12 months  
**LSC:** NR  
**Equation:** NR  
**Precision error:** NR  
**Number of technicians:** NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ishijima (2009)</td>
<td><strong>Original study design:</strong> uncontrolled cohort</td>
<td><strong>Test 1:</strong> sBALP</td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review:</strong> uncontrolled cohort</td>
<td><em>Assay method used:</em> ELISA</td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis:</strong> LS BMD ≤ 70% young adult mean; LS BMD ≤ 80% young adult mean + ≥ 1 fracture</td>
<td><em>Timing of test:</em> baseline; 6 months</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied:</strong> conditions known to influence bone metabolism; renal impairment and/or transplant; patients who had been treated with medication of primary osteoporosis</td>
<td><em>Dietary restrictions:</em> fasting (details NR)</td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given:</strong> no</td>
<td><em>Time of collection:</em> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Treatment:</strong> alendronate 5 mg/day orally for 6 months</td>
<td><em>Storage temperature:</em> NR</td>
</tr>
<tr>
<td></td>
<td><em>N</em> = 45; <em>n</em> with OP = 45; <em>n</em> PMW = 45; <em>n</em> male = 0</td>
<td><em>Delay to freezing:</em> NR</td>
</tr>
<tr>
<td></td>
<td><em>Mean age:</em> 70.2 (SD 7.1) years</td>
<td><em>Time in storage:</em> NR</td>
</tr>
<tr>
<td></td>
<td><em>n</em> with prior fracture: NR</td>
<td><em>Specialist laboratory:</em> yes</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements:</strong> mean LS: g/cm²: 0.7 and T-score: −2.81</td>
<td><em>LSC:</em> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements:</strong> mean sBALP: 28.6 IU/L; uNTX: 57.5 nM BCE/mM</td>
<td><em>Equation:</em> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up:</strong> maximum 6 months</td>
<td><em>Intra-assay CV:</em> maximum: 15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Inter-assay CV:</em> NR</td>
</tr>
</tbody>
</table>

**Test 2:** uNTX
*Assay method used:* ELISA
*Timing of test:* baseline; 6 months
*Sample type:* NR
*Corrected for Cr:* yes
*Dietary restrictions:* NR
*Time of collection:* NR
*Storage temperature:* −70 °C
*Delay to freezing:* NR
*Time in storage:* NR
*Specialist laboratory:* yes
*LSC:* NR
*Equation:* NR
*Intra-assay CV:* maximum: 10%
*Inter-assay CV:* NR
*Number of:
  *samples:* NR
  *replicates per run:* NR
*Analytical sensitivity:* NR
*Upper normal limit:* NR

**Test 3:** DXA
*Area assessed:* LS (L2–L4); units used: kg/cm²
*Timing of test:* baseline; 6 months
*LSC:* NR
*Equation:* NR
*Precision error:* < 1%
*Number of technicians:* NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iwamoto (2004)</td>
<td><strong>Original study design</strong>: uncontrolled cohort</td>
<td><strong>Test 1</strong>: uNTX</td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review</strong>: uncontrolled cohort</td>
<td><strong>Assay method used</strong>: ELISA</td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis</strong>: LS BMD ≤ 70% young adult mean; LS BMD ≤ 80% young adult mean + ≥ 1 fracture</td>
<td><strong>Timing of test</strong>: baseline; 6 and 12 months</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied</strong>: conditions known to influence bone metabolism; medications known to affect bone metabolism</td>
<td><strong>Sample type</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given</strong>: everyone</td>
<td><strong>Corrected for Cr</strong>: yes</td>
</tr>
<tr>
<td></td>
<td><strong>Treatment</strong>: alendronate 5 mg/day orally for 12 months</td>
<td><strong>Dietary restrictions</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>N</strong> = 85; <strong>n with OP</strong> = 85; <strong>n PMW</strong> = 85; <strong>n male</strong> = 0</td>
<td><strong>Time of collection</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Mean age</strong>: 72.2 (SD 7.8; range 55 to 88) years</td>
<td><strong>Storage temperature</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements</strong>: mean LS: g/cm²: 0.574</td>
<td><strong>Delay to freezing</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements</strong>: uNTX: 71.4 nM BCE/mM</td>
<td><strong>Time in storage</strong>: NR</td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up</strong>: NR</td>
<td><strong>Specialist laboratory</strong>: NR</td>
</tr>
<tr>
<td>Iwamoto (2005)</td>
<td><strong>Original study design</strong>: uncontrolled cohort</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review</strong>: uncontrolled cohort</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis</strong>: LS BMD ≤ 70% young adult mean; LS BMD ≤ 80% young adult mean + ≥ 1 fracture</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied</strong>: conditions known to influence bone metabolism; medications known to affect bone metabolism</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given</strong>: no (all patients instructed to take at least 800 mg Ca via diet)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Treatment</strong>: alendronate 5 mg/day orally for 1 year</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>N</strong> = 132; <strong>n with OP</strong> = 132; <strong>n PMW</strong> = 132; <strong>n male</strong> = 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Mean age</strong>: 71.9 (SD 7.5; range 54 to 88) years</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements</strong>: mean LS: g/cm²: 0.576</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements</strong>: uNTX: 68.8 nM BCE/mM</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up</strong>: maximum 12 months</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Population and treatment details</td>
<td>Intervention/test details</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **Kim (2005)**        | **Original study design**: controlled cohort                           | **Test 1**: uNTX  
  **Assay method used**: ELISA  
  **Sample type**: second morning void; Corrected for Cr; yes  
  **Timing of test**: baseline; 3 and 6 months  
  **Dietary restrictions**: NR  
  **Storage temperature**: −20 °C  
  **Delay to freezing**: NR  
  **Time in storage**: NR  
  **Specialist laboratory**: NR  
  **LSC**: NR  
  **Equation**: NR  
  **Intra-assay CV**: NR  
  **Inter-assay CV**: NR  
  **Analytical sensitivity**: NR  
  **Upper normal limit**: NR |
|                       | **Study design as used in this review**: derived single-arm cohort(s)  |                                                                                           |
|                       | **Definition of osteoporosis**: T-score ≤ 2.5 SD below normal mean for Korean PMW at LS |                                                                                           |
|                       | **Exclusion criteria applied**: none reported                          |                                                                                           |
|                       | **Supplemental Ca or vitamin D given**: everyone                      |                                                                                           |
|                       | **Treatment**: alendronate 10 mg/day orally for 1 year  
  N = 50; n with OP = 50; n PMW = 50; n male = 0  
  **Mean age**: 60.3 (SD 8.0) years  
  n with prior fracture: NR  
  **Baseline BMD measurements**: mean LS g/cm²: 0.761 and T-score: −2.99; FN: g/cm²: 0.674 and T-score: −1.85  
  **Baseline BM measurements**: mean uNTX: 111.2 nM BCE/mM |                                                                                           |
|                       | **Follow-up**: maximum 12 months                                       |                                                                                           |
| **Kitatani (2003)**   | **Original study design**: RCT                                         | **Test 2**: DXA  
  **Area assessed**: FN; LS (L1–L4); units used: g/cm²  
  **Timing of test**: baseline; 12 months  
  **LSC**: 4.02%  
  **Equation**: (1.96 × sqrt2) × precision error  
  **Precision error**: 1.8% LS; 1.9% FN  
  **Number of technicians**: 1 |
|                       | **Study design as used in this review**: derived single-arm cohort(s)  |                                                                                           |
|                       | **Definition of osteoporosis**: LS BMD ≤ 70% young adult mean          |                                                                                           |
|                       | **Exclusion criteria applied**: conditions known to influence bone metabolism; glucocorticosteroids; medications known to affect bone metabolism; scoliosis and/or other severe spinal disorders; history of bisphosphonate treatment |                                                                                           |
|                       | **Supplemental Ca or vitamin D given**: only controls                 |                                                                                           |
|                       | **Treatment 1**: etidronate (2 weeks with drug followed by 10 weeks without) 200 mg/day orally for 98 weeks  
  N = 32; n with OP = 32; n PMW = 32; n male = 0  
  **Mean age**: 63.3 (SD 7.4) years  
  n with prior fracture: NR  
  **Baseline BMD measurements**: mean LS g/cm²: 0.679  
  **Baseline BM measurements**: sBALP: 24.8 U/l |                                                                                           |
|                       | **Treatment 2**: etidronate (2 weeks with drug followed by 10 weeks without) 400 mg/day orally for 98 weeks  
  N = 31; n with OP = 31; n PMW = 31; n male = 0  
  **Mean age**: 64.8 (SD 5.6) years  
  n with prior fracture: NR  
  **Baseline BMD measurements**: mean LS g/cm²: 0.684  
  **Baseline BM measurements**: sBALP: 23.8 U/l |                                                                                           |
<p>|                       | <strong>Follow-up</strong>: Range &lt; 6 to 24 months                                  |                                                                                           |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kung (2009)</strong>, 133 Asia English Study dates: NR Data from manufacturer's trial database; full published paper</td>
<td><strong>Original study design</strong>: RCT <strong>Study design as used in this review</strong>: RCT <strong>Definition of osteoporosis</strong>: NR <strong>Exclusion criteria applied</strong>: conditions known to influence bone metabolism; hypersensitivity to bisphosphonate; medications known to affect bone metabolism; renal impairment and/or transplant <strong>Supplemental Ca or vitamin D given</strong>: NR <strong>Treatment</strong>: ibandronate 150 mg monthly orally for 12 months ( N = 596; , n \text{ with } OP = 596; , n \text{ PMW} = 596; , n \text{ male} = 0 ) ( \text{Mean age}: \text{maximum } 85 \text{ years} ) ( n \text{ with prior fracture: NR} ) <strong>Baseline BMD measurements</strong>: NR <strong>Baseline BM measurements</strong>: NR <strong>Follow-up</strong>: Mean: 12.6 months; Range 11.8 to 14.4</td>
<td><strong>Intervention</strong>: BM feedback (sCTX after 3 months treatment) ( N = 300; , n \text{ with } OP = 300; , n \text{ male} = 0; , n \text{ with prior fracture: 62} ) ( \text{Mean age}: 66.3 (SD 7.5) \text{ years} ) <strong>Baseline BMD</strong>: NR <strong>Baseline BM</strong>: 0.64 ng/ml <strong>Intervention</strong>: no BM feedback ( N = 296; , n \text{ with } OP = 296; , n \text{ male} = 0; , n \text{ with prior fracture: 49} ) <strong>Mean age</strong>: 65.6 (SD 7.4) years <strong>Baseline BMD</strong>: NR <strong>Baseline BM</strong>: NR <strong>Test</strong>: sCTX <strong>Assay method used</strong>: NR <strong>Timing of test</strong>: baseline; 3 and 6 months <strong>Dietary restrictions</strong>: overnight fasting <strong>Time of collection</strong>: morning (08.00 to 10.00) <strong>Storage temperature</strong>: NR <strong>Delay to freezing</strong>: NR <strong>Time in storage</strong>: NR <strong>Specialist laboratory</strong>: no <strong>LSC</strong>: &gt; 45% if new user; &gt; 15% if current BP user <strong>Equation</strong>: NR <strong>Intra-assay CV</strong>: NR <strong>Inter-assay CV</strong>: NR <strong>Number of samples replicates per run</strong>: NR <strong>Analytical sensitivity</strong>: NR <strong>Upper normal limit</strong>: NR <strong>Test 1</strong>: sBALP-I <strong>Assay method used</strong>: IRMA <strong>Timing of test</strong>: baseline; 3 months <strong>Dietary restrictions</strong>: NR <strong>Time of collection</strong>: NR <strong>Storage temperature</strong>: −20 °C <strong>Delay to freezing</strong>: NR <strong>Time in storage</strong>: NR <strong>Specialist laboratory</strong>: NR <strong>LSC</strong>: 30 <strong>Equation</strong>: 2.77(CVa^2 + CVi^2)^1/2 <strong>Intra-assay CV</strong>: 5.4% <strong>Inter-assay CV</strong>: 9.6% <strong>Number of samples replicates per run</strong>: NR <strong>Analytical sensitivity</strong>: NR <strong>Upper normal limit</strong>: NR <strong>Test 2</strong>: sBALP-E <strong>Assay method used</strong>: immunocapture enzymatic assay <strong>Timing of test</strong>: baseline; 3 months <strong>Dietary restrictions</strong>: NR <strong>Time of collection</strong>: NR <strong>Storage temperature</strong>: −20 °C <strong>Delay to freezing</strong>: NR</td>
</tr>
</tbody>
</table>
Original study design: uncontrolled cohort

Study design as used in this review: uncontrolled cohort

Definition of osteoporosis: T-score at LS/FN ≤ −2.5

Exclusion criteria applied: none reported

Supplemental Ca or vitamin D given: patients were given Ca depending upon their dietary intake

Treatment: alendronate 10 mg/day orally for 1 year
N = 30; n with OP = 30; n PMW = 30; n male = 0
Mean age: range 52 to 82 years
n with prior fracture: NR
Baseline BMD measurements: NR
Baseline BM measurements: sCTX: 2052 pmol/l;
uNTX: 51.6 BCE/mM

Follow-up: NR

Test 3: DXA
Area assessed: FN; spine (unspecified); units used: NR
Timing of test: baseline; 12 months
LSC: NR
Equation: NR
Precision error: 1% at the LS; 2% at the FN
Number of technicians: two

Test 1: sCTX
Assay method used: ELISA
Timing of test: baseline; 3 and 6 months
Dietary restrictions: NR
Time of collection: NR
Storage temperature: −20 °C
Delay to freezing: NR
Time in storage: NR
Specialist laboratory: NR
LSC: 53.6;
Equation: 2.77(CVa^2 + CVi^2)^1/2
Intra-assay CV: range: 5% to 7%
Inter-assay CV: maximum 10%
Number of:
samples: NR
replicates per run: NR
Analytical sensitivity: NR
Upper normal limit: NR

Test 2: uNTX
Assay method used: ELISA
Timing of test: baseline; 3 and 6 months
Sample type: second morning void
Corrected for Cr: yes
Dietary restrictions: NR
Storage temperature: −70 °C
Delay to freezing: NR
Time in storage: NR
Specialist laboratory: NR
LSC: 40.6%
Equation: 2.77(CVa^2 + CVi^2)^1/2
Intra-assay CV: 5.0%
Inter-assay CV: maximum 10%
Number of:
samples: NR
replicates per run: NR
Analytical sensitivity: NR
Upper normal limit: NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| **Lane (2000)**, USA/Canada | **Test 3**: DXA  
Area assessed: FN; LS (L2–L4);  
units used: NR  
Timing of test: baseline; 12 months  
LSC: NR  
Equation: $2.77(CV_a^2 + CV_i^2)^{1/2}$  
Precision error: 1% at the LS; 2% at the FN  
Number of technicians: two | **Test 1**: sBALP  
Assay method used: EIA  
Timing of test: baseline; 1, 3, 6, 9, 18 and 24 months  
Dietary restrictions: NR  
Time of collection: morning (between 10:00 and 11:00 hours)  
Storage temperature: NR  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: NR  
LSC: NR  
Equation: $t \times \sqrt{2 \times \text{median long-term CV}}$ for each measure  
Precision error: NR  
Number of:  
samples: NR  
replicates per run: NR  
Analytical sensitivity: NR  
Upper normal limit: NR |
| **Masaryk (2002)**, eastern Europe, Slovak | **Test 2**: DXA  
Area assessed: FN; hip (unspecified);  
LS (unspecified);  
units used: g/cm²  
Timing of test: baseline; 6, 12, 18 and 24 months  
LSC: NR  
Equation: $t \times \sqrt{2 \times \text{median long-term CV}}$ for each measure  
Precision error: NR  
Number of:  
samples: NR  
replicates per run: NR  
Analytical sensitivity: NR  
Upper normal limit: NR | **Test 1**: uNTX  
Assay method used: EIA  
Timing of test: baseline; 3 months  
Sample type: 2 hour  
Corrected for Cr: yes  
Dietary restrictions: unclear  
Time of collection: NR  
Specialist laboratory: unclear  
LSC: NR  
Equation: $t \times \sqrt{2 \times \text{median long-term CV}}$ for each measure  
Precision error: NR  
Number of:  
samples: NR  
replicates per run: NR  
Analytical sensitivity: NR  
Upper normal limit: NR |

**Original study design**: RCT  
**Study design as used in this review**: derived single-arm cohort(s)  
**Definition of osteoporosis**: T-score at FN $\leq -2.5$; T-score at LS/hip $\leq -2.5$  
**Exclusion criteria applied**: abnormalities on spinal radiographs – precluded lumbar QCT or DXA; liver dysfunction; Renal impairment and/or transplant; secondary osteoporosis other than for rheumatic diseases  
**Supplemental Ca or vitamin D given**: only those with deficiency  
**Treatment**: Teriparatide 40 µg/day SC (duration NR)  
$N = 28$; $n$ with OP = 28; $n$ PMW = 28; $n$ male = 0  
Mean age: NR  
$n$ with prior fracture: NR  
Baseline BMD measurements: mean LS: g/cm²: 0.84; hip: g/cm²: 0.70  
Baseline BM measurements: mean sBALP: 14.0 U/L  
**Follow-up**: minimum 1 year

**Original study design**: uncontrolled cohort  
**Study design as used in this review**: uncontrolled cohort  
**Definition of osteoporosis**: T-score at LS/hip $\leq -2.5$  
**Exclusion criteria applied**: none reported  
**Supplemental Ca or vitamin D given**: everyone  
**Treatment**: alendronate 10 mg/day orally for 12 months  
$N = 50$; $n$ with OP = 50; $n$ PMW = 50; $n$ male = 50  
Mean age: 64.2 (SD 7.07; range 49 to 78) years  
$n$ with prior fracture: NR  
Baseline BMD measurements: mean LS T-score: $-2.62$; FN T-score: $-2.52$  
Baseline BM measurements: mean uNTX: 68.51 unclear (units not reported)  
**Follow-up**: maximum 12 months
### Miller (2008), multinational study

**Study design:** controlled cohort

**Study dates:** NR

**Full published paper:**

#### Population and treatment details

**Definition of osteoporosis:** T-score at LS/hip $\leq -2.5 + \geq 1$

**Exclusion criteria applied:** glucocorticosteroids; medications known to affect bone metabolism; antiresorptive treatment other than alendronate or risidronate

**Supplemental Ca or vitamin D given:** NR

**Treatment:** teriparatide 20 µg/day SC/IM daily for 12 months

$N = 317; n$ with OP $= 317; n$ male $= 0$

**Mean age:** NR

**Baseline BMD measurements:** NR

**Baseline BM measurements:** NR

**Follow-up:** NR

### Moro-Alvarez (2010), western Europe study

**Study design:** uncontrolled cohort

**Study dates:** NR

**Abstract:**

#### Population and treatment details

**Definition of osteoporosis:** NR

**Exclusion criteria applied:** NR

**Supplemental Ca or vitamin D given:** NR

**Treatment:** strontium ranelate 2 g/day orally for 12 to 24 months

$N = 66; n$ with OP $= 66; n$ PMW $= 66; n$ male $= 0$

**Mean age:** 68.0 (range 51 to 87) years

**N with prior fracture:** NR

**Baseline BMD measurements:** NR

**Baseline BM measurements:** NR

**Follow-up:** maximum 24 months

---

### Test 1: sP1NP

**Assay method used:** electrochemiluminescence

**Timing of test:** baseline; 0.5, 1, 2, 3, 4, 5, 6, 9 and 12 months

**Dietary restrictions:** NR

**Time in storage:** NR

**Specialist laboratory:** yes

**LSC:** NR

**Equation:** NR

**Intra-assay CV:** NR

**Inter-assay CV:** NR

**Analytical sensitivity:** NR

**Upper normal limit:** NR

### Test 2: sCTX

**Assay method used:** electrochemiluminescence

**Timing of test:** baseline; 12 and 24 months

**Dietary restrictions:** NR

**Time in storage:** NR

**Specialist laboratory:** NR

**LSC:** NR

**Equation:** NR

**Intra-assay CV:** NR

**Inter-assay CV:** NR

**Number of:**

- samples: NR

- replicates per run: NR

**Analytical sensitivity:** NR

**Upper normal limit:** NR

---

### Test 2: DXA

**Area assessed:** FN; LS (L2–L4); total body; trochanter; units used: NR

**Timing of test:** baseline; 12 months

**LSC:** NR

**Equation:** NR

**Precision error:** NR

**Number of technicians:** NR
Reginster (2004),132

Study design: post hoc subgroup analysis of a RCT

Original study design: post hoc subgroup analysis of a RCT

Test 1: DXA
Area assessed: FN; LS (L2–L4); total hip; units used: NR
Timing of test: baseline; 24 months
LSC: NR
Equation: NR
Precision error: 1.2%
Number of technicians: NR

Test 1: sBALP
Assay method used: IRMA
Timing of test: baseline; 6, 12, 24 and 36 months
Dietary restrictions: NR
Time of collection: NR
Storage temperature: NR
Delay to freezing: NR
Time in storage: NR
Specialist laboratory: NR
LSC: NR
Equation: NR
Intra-assay CV: maximum: 8%
Inter-assay CV: maximum: 7%
Number of replicates per run: NR
Analytical sensitivity: 2.6 ng/ml
Upper normal limit: NR

Test 1: sP1NP
Assay method used: RIA
Timing of test: baseline; 6, 12, 24 and 36 months
Dietary restrictions: NR
Time of collection: NR
Storage temperature: NR
Delay to freezing: NR
Time in storage: NR
Specialist laboratory: NR
LSC: NR
Equation: NR
Intra-assay CV: NR
Inter-assay CV: NR
Number of samples: NR
replicates per run: NR
Analytical sensitivity: NR
Upper normal limit: NR

Test 3: uCTX
Assay method used: ELISA
Timing of test: baseline; 6, 12, 24 and 36 months
Sample type: NR
### Reyes-Garcia (2010),58

**Original study design:** uncontrolled cohort  
**Study dates:** NR  
**Full published paper**

<table>
<thead>
<tr>
<th>Study Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Corrected for Cr:</strong> yes &lt;br&gt; <strong>Dietary restrictions:</strong> NR &lt;br&gt; <strong>Time of collection:</strong> NR &lt;br&gt; <strong>Storage temperature:</strong> NR &lt;br&gt; <strong>Delay to freezing:</strong> NR &lt;br&gt; <strong>Time in storage:</strong> NR &lt;br&gt; <strong>Specialist laboratory:</strong> NR &lt;br&gt; <strong>LSC:</strong> NR &lt;br&gt; <strong>Equation:</strong> NR &lt;br&gt; <strong>Intra-assay CV:</strong> NR &lt;br&gt; <strong>Inter-assay CV:</strong> NR &lt;br&gt; <strong>Number of samples:</strong> NR &lt;br&gt; <strong>Number of replicates per run:</strong> NR &lt;br&gt; <strong>Analytical sensitivity:</strong> NR &lt;br&gt; <strong>Upper normal limit:</strong> NR</td>
<td></td>
</tr>
</tbody>
</table>

**Test 1:** sBALP  
*Assay method used:* ELISA  
*Timing of test:* baseline; 3, 6 and 12 months  
*Dietary restrictions:* overnight fasting; <br> *Time of collection:* morning <br> *Storage temperature:* –80 °C <br> *Delay to freezing:* NR <br> *Time in storage:* NR <br> *Specialist laboratory:* NR <br> *LSC:* NR <br> *Equation:* NR  
*Intra-assay CV:* 4.2% <br> *Inter-assay CV:* 7.2% <br> *Analytical sensitivity:* NR <br> *Upper normal limit:* NR

**Test 2:** sCTX  
*Assay method used:* electrochemiluminescence  
*Timing of test:* baseline; 3, 6 and 12 months  
*Dietary restrictions:* overnight fasting; <br> *Time of collection:* morning <br> *Storage temperature:* –80 °C <br> *Delay to freezing:* NR <br> *Time in storage:* NR <br> *Specialist laboratory:* NR <br> *LSC:* NR <br> *Equation:* NR  
*Intra-assay CV:* 4.2% <br> *Inter-assay CV:* 5.1% <br> *Analytical sensitivity:* NR <br> *Upper normal limit:* NR

**Test 3:** DXA  
*Area assessed:* FN; LS (L2–L4); units used: g/cm²  
*Timing of test:* baseline; 12 months  
*LSC:* NR  
*Equation:* NR  
*Precision error:* < 1%  
*Number of technicians:* NR

---

**Test 1:** sBALP  
*Assay method used:* ELISA  
*Timing of test:* baseline; 3, 6 and 12 months  
*Dietary restrictions:* overnight fasting; <br> *Time of collection:* morning <br> *Storage temperature:* –80 °C <br> *Delay to freezing:* NR <br> *Time in storage:* NR <br> *Specialist laboratory:* NR <br> *LSC:* NR  
*Equation:* NR  
*Intra-assay CV:* 4.2%  
*Inter-assay CV:* 7.2%  
*Analytical sensitivity:* NR  
*Upper normal limit:* NR

**Test 2:** sCTX  
*Assay method used:* electrochemiluminescence  
*Timing of test:* baseline; 3, 6 and 12 months  
*Dietary restrictions:* overnight fasting; <br> *Time of collection:* morning <br> *Storage temperature:* –80 °C <br> *Delay to freezing:* NR <br> *Time in storage:* NR <br> *Specialist laboratory:* NR <br> *LSC:* NR  
*Equation:* NR  
*Intra-assay CV:* 4.2%  
*Inter-assay CV:* 7.2%  
*Analytical sensitivity:* NR  
*Upper normal limit:* NR

**Test 3:** DXA  
*Area assessed:* FN; LS (L2–L4); units used: g/cm²  
*Timing of test:* baseline; 12 months  
*LSC:* NR  
*Equation:* NR  
*Precision error:* < 1%  
*Number of technicians:* NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Roche (2007)</strong></td>
<td><strong>Original study design:</strong> RCT</td>
<td><strong>Intervention:</strong> BM feedback (sCTX)</td>
</tr>
<tr>
<td></td>
<td><strong>Study design as used in this review:</strong> RCT</td>
<td><strong>N = NR; n with OP = NR;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Definition of osteoporosis:</strong> NR</td>
<td><strong>n male = 0; n with prior fracture:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion criteria applied:</strong> history major upper GI disease; hypersensitivity to bisphosphonate</td>
<td><strong>Mean age:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Supplemental Ca or vitamin D given:</strong> NR</td>
<td><strong>Baseline BMD:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Treatment:</strong> ibandronate 150 mg monthly orally for 6 months</td>
<td><strong>Baseline BM:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>N = 781; n with OP = 781; n PMW = 781; n male = 0</strong></td>
<td><strong>Intervention:</strong> no BM feedback</td>
</tr>
<tr>
<td></td>
<td><strong>Mean age:</strong> NR</td>
<td><strong>N = NR; n with OP = NR;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>n with prior fracture:</strong> NR</td>
<td><strong>n male = 0; n with prior fracture:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BMD measurements:</strong> NR</td>
<td><strong>Mean age:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Baseline BM measurements:</strong> NR</td>
<td><strong>Baseline BMD:</strong> NR</td>
</tr>
<tr>
<td></td>
<td><strong>Follow-up:</strong> maximum 6 months</td>
<td><strong>Baseline BM:</strong> NR</td>
</tr>
</tbody>
</table>

| **Roche (2009)**  | **Original study design:** RCT                                                                    | **Test:** sCTX                                                                            |
|                   | **Study design as used in this review:** RCT                                                       | **Assay method used:** NR                                                                 |
|                   | **Definition of osteoporosis:** NR                                                                | **Timing of test:** baseline; 3 months in the feedback arm; 6 months                      |
|                   | **Exclusion criteria applied:** conditions known to influence bone metabolism; history major upper GI disease; hypersensitivity to bisphosphonate; medications known to affect bone metabolism; specific prior treatment – give details; bisphosphonate treatment within prior 6 months | **Dietary restrictions:** NR                                                              |
|                   | **Supplemental Ca or vitamin D given:** NR                                                        | **Time of collection:** NR                                                                |
|                   | **Treatment:** ibandronate 150 mg monthly orally for 6 months                                     | **Storage temperature:** NR                                                               |
|                   | **N = 585; n with OP = 585; n PMW = 585; n male = 0**                                              | **Delay to freezing:** NR                                                                |
|                   | **Age range:** 55 to 85                                                                           | **Time in storage:** NR                                                                  |
|                   | **n with prior fracture:** NR                                                                    | **Specialist laboratory:** NR                                                             |
|                   | **Baseline BMD measurements:** NR                                                                 | **LSC:** 30                                                                              |
|                   | **Baseline BM measurements:** NR                                                                  | **Equation:** NR                                                                         |
|                   | **Follow-up:** maximum 6 months                                                                   | **Intra-assay CV:** NR                                                                   |
|                   |                                                                                                  | **Inter-assay CV:** NR                                                                   |
|                   |                                                                                                  | **Number of:**                                                                          |
|                   |                                                                                                  | **samples:** NR                                                                          |
|                   |                                                                                                  | **replicates per run:** NR                                                               |
|                   |                                                                                                  | **Analytical sensitivity:** NR                                                           |
|                   |                                                                                                  | **Upper normal limit:** NR                                                               |
### Study Details

<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| **Roche (2009),** Roche (2009), 149 western Europe | **Original study design:** RCT | **Intervention:** BM feedback (sCTX after 5 week)  
**Number of:**  
- samples: NR  
- replicates per run: NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR  
**Intervention:** BM feedback (sCTX after 5 week)  
**Number of:**  
- samples: NR  
- replicates per run: NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR  |
| **Study design as used in this review:** RCT | **Treatment:** ibandronate 150 mg monthly orally for 12 months  
- N = 596; n with OP = 596; n PMW = 596; n male = 0  
**Mean age:** NR  
**n with prior fracture:** NR  
**Baseline BMD measurements:** NR  
**Baseline BM measurements:** NR | **Follow-up:** minimum 12 months |
| **Definition of osteoporosis:** NR | | |
| **Exclusion criteria applied:** < 55 years old; history major upper GI disease; hypersensitivity to bisphosphonate; medications known to affect bone metabolism; bisphosphonates within last 6 months | | |
| **Supplemental Ca or vitamin D given:** NR | | |
| **Test:** sCTX | | |
| **Assay method used:** IRMA  
**Timing of test:** baseline; 5 weeks; 3, 6 and 12 months  
**Dietary restrictions:** NR  
**Time of collection:** NR  
**Storage temperature:** NR  
**Delay to freezing:** NR  
**Time in storage:** NR  
**Specialist laboratory:** NR  
**LSC:** NR  
**Equation:** NR  
**Intra-assay CV:** NR  
**Inter-assay CV:** NR  
**Number of:**  
- samples: NR  
- replicates per run: NR  
**Analytical sensitivity:** NR  
**Upper normal limit:** NR | **Test 1:** sBALP  
**Area assessed:** FN; LS (L2–L4); units used: g/cm²  
**Timing of test:** annually |
| **Study design as used in this review:** derived single-arm cohort(s) | | |
| **Definition of osteoporosis:** at least two vertebral fractures; T-score at L5/hip ≤ –2.5 | | |
| **Exclusion criteria applied:** conditions known to influence bone metabolism; lifestyle known to influence bone metabolism; liver dysfunction; medications known to affect bone metabolism; renal impairment and/or transplant | | |
| **Supplemental Ca or vitamin D given:** everyone | | |
| **Treatment:** raloxifene 60 or 120 mg/day orally (duration NR)  
- N = 1650; n with OP = 1650; n PMW = 1650; n male = 0  
**Mean age:** 67.3 (SD 6.73) years  
**n with prior fracture:** vertebral: 626  
**Baseline BMD measurements:** mean FN: g/cm²: 0.62; LS: g/cm²: 0.83  
**Baseline BM measurements:** mean sBALP: 16.36 µg/l | | |
<p>| <strong>Follow-up:</strong> maximum 3 years | | |
| <strong>Sarkar (2004),</strong> Sarkar (2004), 164 multinational | <strong>Original study design:</strong> post hoc analysis of a RCT | | |
| <strong>Study design as used in this review:</strong> derived single-arm cohort(s) | | |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
<th>LSC: NR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study design as used in this review: uncontrolled cohort</td>
<td>Precision error: NR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Definition of osteoporosis: LS BMD ≤ 70% young adult mean; LS BMD ≤ 80% young adult mean + ≥ 1 fracture</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exclusion criteria applied: conditions known to influence bone metabolism; medications known to affect bone metabolism; previous treatment with bisphosphonates</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supplemental Ca or vitamin D given: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treatment: alendronate 5 mg/day or 35 mg/week, or risedronate 2.5 mg/day or 17.5 mg/week orally for mean 3.2 years ± 2.0 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = 251; n with OP = 251; n PMW = 251; n male = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean age: 70.5 (SD 8.9) years</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n with prior fracture: any: 154; vertebral: 144; non-vertebral: 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline BMD measurements: mean LS: g/cm²: 0.77 Baseline BM measurements:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow-up: Mean: 3.2 years; Range 1 to 8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddiqi (2010), 106 UK English Study dates: NR</td>
<td>Original study design: uncontrolled cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Study design as used in this review: uncontrolled cohort</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Definition of osteoporosis: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exclusion criteria applied: glucocorticosteroids</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Supplemental Ca or vitamin D given: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test 1: sBALP Assay method used: EIA Timing of test: baseline; 6-month intervals; study end</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test 2: uNTX Assay method used: ELISA Timing of test: baseline; 6-month intervals; study end Sample type: second morning void Corrected for Cr: yes Dietary restrictions: none Storage temperature: NR Delay to freezing: NR Time in storage: NR Specialist laboratory: unclear LSC: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test 3: DXA Area assessed: LS (unspecified); units used: g/cm² Timing of test: baseline; every 6 months LSC: NR Equation: NR Precision error: 0.5% (SD 0.5) Number of technicians: NR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Study Population and treatment details

**Treatment**: teriparatide for 18 months

- N = 28; n with OP = 28; n PMW = NR; n male = 0
- Mean age: 74 (range 50 to 85) years
- n with prior fracture: NR

**Baseline BMD measurements**: spine: g/cm²: 0.787
**Baseline BM measurements**: mean sP1NP: 28 µg/l

**Follow-up**: NR

---

**Test 2**: DXA

- Area assessed: spine (unspecified); units used: NR
- Timing of test: baseline; 18 months
- LSC: NR
- Equation: NR
- Precision error: NR
- Number of: NR
- Analytical sensitivity: NR
- Upper normal limit: NR

**Test 1**: sP1NP

- Assay method used: NR
- Timing of test: baseline; 1, 3, 6, 12 and 24 months
- Dietary restrictions: NR
- Time of collection: NR
- Storage temperature: NR
- Delay to freezing: NR
- Time in storage: NR
- Specialist laboratory: NR
- LSC: NR
- Equation: NR
- Intra-assay CV: NR
- Inter-assay CV: NR
- Number of: NR
- Analytical sensitivity: NR
- Upper normal limit: NR

**Test 3**: biopsy

- Site: iliac crest
- Number: NR
- Needle: NR
- Technique: NR
- Embedding method: NR
- Anaesthesia: NR
- Number of clinicians: NR

---

**Steepan (2008)**, multinational English

**Study design as used in this review**: uncontrolled cohort

- **Exclusion criteria applied**: none reported;
- **Supplemental Ca or vitamin D given**: NR

**Treatment**: teriparatide 20 µg/day SC for 24 months

- N = 66; n with OP = 66; n PMW = 66; n male = 0
- Mean age: 68.0 years
- n with prior fracture: 41
- Baseline BMD measurements: mean LS T-score: –2.8; hip T-score: –1.7
- Baseline BM measurements: NR

**Follow-up**: maximum 24 months
<table>
<thead>
<tr>
<th>Study</th>
<th>Population and treatment details</th>
<th>Intervention/test details</th>
</tr>
</thead>
</table>
| Tsujimoto (2011), 146 Asia English | **Original study design**: RCT  
**Study design as used in this review**: derived single-arm cohort(s)  
**Definition of osteoporosis**: BMD LS (L2–L4) < 65% of young adult mean and age ≥ 55; BMD LS (L2–L4) < 70% of young adult mean and age ≥ 65; LS BMD ≤ 80% young adult mean + ≥ 1 fracture  
**Exclusion criteria applied**: conditions known to influence bone metabolism; lifestyle known to influence bone metabolism; medications known to affect bone metabolism; bisphosphonate or raloxifene in last 3 months  
**Supplemental Ca or vitamin D given**: everyone  
**Treatment**: teriparatide 20 µg/day SC for 12 months  
N = 136; n with OP = 136; n PMW = 127; n male = 9  
Mean age: 69.2 (SD 6.3) years  
n with prior fracture: vertebral: 54  
Baseline BMD measurements: mean LS: g/cm²: 0.639  
Baseline BM measurements: mean CTX: 0.54 µg/ml; sP1NP: 15.53 ng/ml; sBALP: 55.7 ng/ml  
**Follow-up**: NR |  
**Test 1**: sBALP  
Assay method used: ostease assay (type NR)  
Timing of test: baseline; 1, 3, 6 and 12 months  
Dietary restrictions: overnight/morning fasting  
Time of collection: morning  
Storage temperature: −20 °C or −70 °C  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: unclear  
LSC: NR  
Equation: NR  
Intra-assay CV: maximum: 4.4%  
Inter-assay CV: maximum: 7.3%  
Number of samples: NR  
replicates per run: NR  
Analytical sensitivity: NR  
Upper normal limit: NR  
**Test 2**: sP1NP  
Assay method used: RIA  
Timing of test: baseline; 1, 3, 6 and 12 months  
Dietary restrictions: overnight/morning fasting;  
Time of collection: morning  
Storage temperature: −20 °C or −70 °C  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: unclear  
LSC: > 10 µg/l  
Equation: NR  
Intra-assay CV: range: 1.7% to 2.9%  
Inter-assay CV: range 3% to 6%  
Number of samples: NR  
replicates per run: NR  
Analytical sensitivity: NR  
Upper normal limit: NR  
**Test 3**: sCTX  
Assay method used: ELISA  
Timing of test: baseline; 1, 3, 6 and 12 months  
Dietary restrictions: overnight/morning fasting  
Time of collection: morning  
Storage temperature: −20 °C or −70 °C  
Delay to freezing: NR  
Time in storage: NR  
Specialist laboratory: unclear  
LSC: NR  
Equation: NR  
Intra-assay CV: range: 4.9% to 6.4%  
Inter-assay CV: range 5% to 5.1%  
Number of samples: NR  
replicates per run: NR |
Study Population and treatment details

Intervention/test details

Original study design: post hoc subgroup analysis of a RCT

Study design as used in this review: derived single-arm cohort(s)

Definition of osteoporosis: T-score at LS/hip ≤ –2.5

Exclusion criteria applied: abnormalities on spinal radiographs – precluded lumbar QCT or DXA; conditions known to influence bone metabolism; history hip fracture; medications known to affect bone metabolism (none of patients received any bone-active medications)

Supplemental Ca or vitamin D given: Ca, everyone; vitamin D, NR

Treatment: alendronate 10 mg/day orally for at least 1 year
N = 180; n with OP = 180; n PMW = 180; n male = 0
Mean age: NR
n with prior fracture: NR
Baseline BMD measurements: NR
Baseline BM measurements: NR

Follow-up: NR

Test 1: sBAP

Assay method used: EIA
Timing of test: baseline; 3, 6 and 12 months
Dietary restrictions: NR
Time of collection: NR
Storage temperature: –70 °C
Delay to freezing: NR
Time in storage: NR
Specialist laboratory: yes
LSC: –20%
Equation: \( t \sqrt{2 \times \text{median long-term intraindividual variability of women in the placebo group}} \)
Intra-assay CV: 2.9%
Inter-assay CV: range 5.8% to 9.3%
Number of:
samples: NR
replicates per run: NR
Analytical sensitivity: NR
Upper normal limit: NR

Test 2: DXA

Area assessed: FN; LS (L1–L4); total body; units used: NR
Timing of test: baseline; 3 months; 6 months; 12 months; 18 months; 24 months; 36 months
LSC: 3.88% (LS); 5.04% FN
Equation: \( t \times \sqrt{2 \times \text{median long-term intra-individual variability of women in the placebo group}} \)
Precision error: NR
Number of technicians: NR

BCE, bone collagen equivalents; BM, bone turnover marker; BP, bisphosphonate; Ca, calcium; Cr, creatinine; ECL, electrochemiluminescence; EIA, enzyme immunoassay; FN, femoral neck; GCS, glucocorticoid steroid; IM, intramuscular; intact P1NP, measurement of the trimetric forms only; i.v., intravenous; LLOD, lowest level of detection; LLOQ, lowest limit of quantification; LS, lumbar spine; LSC, least significant change; NR, not reported; OP, osteoporosis; PMW, postmenopausal women; QCT, quantitative computed tomography; RIA, radioimmunoassay; SC, subcutaneous; SD, standard deviation; SE, standard error; sqrt, square root; TB, total body; total P1NP, measurement of the mono- and trimetric forms.
Appendix 5  Summary of the methods for modelling adherence and treatment management

<table>
<thead>
<tr>
<th>Study</th>
<th>Question</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charpurlat</td>
<td>Country?</td>
<td>USA</td>
</tr>
<tr>
<td>(2002)(^{167})</td>
<td>Type of model?</td>
<td>Decision tree (3 months) and Markov model (post 3 months to 5 years)</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>To explore the potential value of bone markers to monitor antiresorptive treatments of osteoporosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two different treatment pathways were compared: (1) without specific follow-up (no BM, no BMD, only simple short-term follow-up to rule out adverse reactions) and (2) follow-up including an early measurement of a serum marker of bone resorption (3 months after beginning treatment for post-menopausal osteoporosis)</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>The terms adherence and compliance appeared to be used interchangeably. Compliance with treatment was considered as a dichotomous variable: patients were assumed to take 100% of their medication or not to take it at all</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>Adherence rates were assumed to be constant over the time horizon of the model. In the upfront decision tree part of the model, the population cohort was divided into adherent and non-adherent groups according to an adherence rate estimate, which was assumed to be 50% in the base case</td>
</tr>
<tr>
<td></td>
<td></td>
<td>It was assumed that bone marker feedback would not increase adherence, owing to lack of evidence. The effect of feedback on adherence was varied in sensitivity analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The adherent population was further divided into responders and non-responders, given a 3% non-response rate obtained from clinical experts. No response was described as a real bone tissue resistance. This was varied from 3% to 30% in sensitivity analysis</td>
</tr>
<tr>
<td></td>
<td>How was treatment management modelled?</td>
<td>Permitted in the model if: ‘the BM measurement leads to the conclusion that compliance or response to treatment was inadequate. The second treatment might pertain to a different pharmacological class, such as parathyroid hormone’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>It was assumed that markers were able to identify true and false responders and non-responders. So no test accuracy data were included in the model</td>
</tr>
<tr>
<td>Earnshaw</td>
<td>Country?</td>
<td>USA</td>
</tr>
<tr>
<td>(2007)(^{183})</td>
<td>Type of model?</td>
<td>A Markov cohort model</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>To model the cost-effectiveness of a monthly and weekly bisphosphonates as an example and explicitly examine differences in costs and outcomes related to persistence</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>Only persistence was included in the model. Persistence was the time spent on medication</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>Persistence rates were derived from a clinical trial. Persistence was found to be 39% at 6 months for weekly alendronate (and 57% for ibandronate). In this model, they assumed persistence on bisphosphonates would be that value. A continual decrease in persistence was modelled post 6 months over 5 years by extrapolating data from persistence studies and longer-term drug utilisation patterns from a UK GP research database. These data approximated a Weibull distribution. Discontinuation is modelled as having no treatment. There was no switch to an alternative</td>
</tr>
</tbody>
</table>
### Study Question Summary

Monthly bisphosphonate use was extrapolated using a 50% increase in persistence compared with weekly medication from the first clinical trial mentioned above.

Sensitivity analyses were conducted on these rates.

**How was treatment management modelled?** Not applicable

---

#### Hiligsmann (2009)\(^{188}\)

<table>
<thead>
<tr>
<th>Country?</th>
<th>Belgium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of model?</td>
<td>Markov microsimulation</td>
</tr>
<tr>
<td>Study objective?</td>
<td>To describe and validate an original Markov microsimulation model to accurately assess the cost-effectiveness of the prevention and treatment of osteoporosis</td>
</tr>
<tr>
<td>Adherence definition?</td>
<td>Adherence was divided into compliance and persistence</td>
</tr>
<tr>
<td>How was adherence modelled?</td>
<td>Medication costs and fracture reduction efficacy were assumed to be proportional to compliance</td>
</tr>
<tr>
<td>Compliance rate</td>
<td>The compliance rate was estimated at 70.5% for persistent women from a clinical study</td>
</tr>
<tr>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

---

#### Hiligsmann (2010)\(^{189}\)

<table>
<thead>
<tr>
<th>Country?</th>
<th>Belgium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of model?</td>
<td>Markov microsimulation</td>
</tr>
<tr>
<td>Study objective?</td>
<td>To evaluate the potential clinical and economic implications of non-adherence to bisphosphonate therapy</td>
</tr>
<tr>
<td>Adherence definition?</td>
<td>ISPOR definition</td>
</tr>
<tr>
<td>How was adherence modelled?</td>
<td>A Markov microsimulation model was developed. Persistence and compliance with bisphosphonate therapies were derived from a large observational (Belgian) study. Two persistence scenarios were modelled: full persistence over 3 years and real-world persistence</td>
</tr>
<tr>
<td>Real-world persistence</td>
<td>Real-world persistence: assumed that 30%, 12%, 18%, and 15% of patients stopped drug therapy after 3 months, 6 months, 1 year, and 2 years from the same clinical study</td>
</tr>
<tr>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>
Study Question Summary

Compliance measured as MPR ranged from 10% to 100% in a data set. There was a gradient of compliance rates. The relative risk (RR) of fracture was dependent on the MPR value and the drug cost was assumed to be proportional to the MPR. It was assumed that the effectiveness of oral bisphosphonates in the meta-analysis was applicable to a population with an MPR value of 80%

For hip fracture, a linear reduction between the MPR value and the probability of fracture was suggested by a Belgian study. A non-linear relationship for non-hip fracture was found from a US observational study.

How was treatment management modelled?
Not applicable

Hiligsmann (2011)190

Country?
Belgium

Type of model?
Markov microsimulation

Study objective?
To estimate the cost-effectiveness of denosumab compared with oral bisphosphonates (branded and generic drugs) in the treatment of post-menopausal osteoporotic women in Belgium

Adherence definition?
Adherence was divided into compliance and persistence

Compliance: the extent to which a patient acts in accordance with the prescribed interval and dose of a dosing regimen. Often expressed as the number of doses taken divided by the number of doses prescribed (MPR)

Persistence: the duration of time from initiation to discontinuation of therapy. The proportion of patients receiving treatment at a given time period

How was adherence modelled?
Adherence to alendronate in a real-life setting was derived from a Belgian study. It was divided into compliance and persistence

Persistence: 42.5% of those who initiated treatment discontinued within 6 months. These were given no treatment benefit but incurred 3 months’ treatment cost. Another 18.1%, 8.3%, 5.6% and 4.1% were dropped off therapy at 1 year, 1.5 years, 2 years and 2.5 years, respectively. Patients discontinuing therapy received no further treatment

Compliance: patients were considered compliant if their MPR was ≥ 80%. The probabilities of being less than 80% compliant were 23.9%, 4% and 1.2% in the first, second and following years of therapy, respectively. From the Belgian study, poor compliance was associated with an increased fracture rate of 35% for hip fractures. From an alternative source, it was assumed that for non-hip fractures the increase in fracture rate would only be 17%. The relative risks of fracture from the meta-analysis were associated with a population with a MPR of ≥ 80%, so if alendronate was assumed to reduce the risk of hip fracture by 38% then non-compliant women would receive a reduction in hip fractures of 16% (0.62 x 1.35 = 0.837). For poorly compliant women, the drug cost was restricted to 80% of full price

Sensitivity analysis was done on reducing poor compliance and discontinuation rates by 25%

How was treatment management modelled?
Not applicable

Hiligsmann (2010)191

Country?
Belgium

Type of model?
Markov microsimulation

Study objective?
To estimate the clinical and economic burden of non-adherence with oral bisphosphonates in osteoporotic patients and the potential cost-effectiveness of adherence-enhancing interventions
<table>
<thead>
<tr>
<th>Study</th>
<th>Question</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>Adherence was divided into compliance and persistence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compliance: the extent to which a patient acts in accordance with the prescribed interval and dose of a dosing regimen. Often expressed as the number of doses taken divided by the number of doses prescribed (MPR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persistence: the duration of time from initiation to discontinuation of therapy. The proportion of patients receiving treatment at a given time period</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>Adherence to alendronate in a real-life setting was derived from a Belgian study. It was divided into compliance and persistence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persistence: 42.5% of those who initiated treatment discontinued within 6 months. These patients were given no treatment benefit but incurred 3 months’ treatment cost. Another 18.1% and 13.9% were dropped off therapy at 1 year and 2 years, respectively. Patients discontinuing therapy received no further treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compliance: patients were considered compliant if their MPR was $\geq 80%$. The probabilities of being less than 80% compliant were 23.9%, 4% and 1.2% in the first, second and following years of therapy, respectively. From the Belgian study, poor compliance was associated with an increased fracture rate of 35% for hip fractures. From an alternative source, it was assumed that for non-hip fractures the increase in fracture rate would only be 17%. The relative risks of fracture from the meta-analysis were associated with a population with a MPR of $\geq 80%$, so if alendronate was assumed to reduce the risk of hip fracture by 38% then non-compliant women would receive a reduction in hip fractures of 16% ($0.62 \times 1.35 = 0.837$). For poorly compliant women, the drug cost was restricted to 80% of full price</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sensitivity analyses were conducted assuming that adherence rates were 10%, 25%, or 50% higher than in the real-world scenario. MPR thresholds of 70% and 90% were examined</td>
</tr>
<tr>
<td></td>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
<tr>
<td></td>
<td>Type of model?</td>
<td>Markov microsimulation</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>To evaluate the impact of all aspects of medication non-adherence on the cost-effectiveness of osteoporosis screening (by DXA)</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>ISPOR definition:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adherence: a general term, encompassing two different constructs, i.e. persistence and compliance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compliance: the extent to which a patient acts in accordance with the prescribed interval and dose of a dosing regimen. Often expressed as the number of doses taken divided by the number of doses prescribed (MPR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persistence: the duration of time from initiation to discontinuation of therapy. The proportion of patients receiving treatment at a given time period</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Primary non-adherence: where patients are diagnosed with osteoporosis but did not take any medication</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>Adherence to alendronate in a real-life setting was derived from a Belgian study. It was divided into compliance, persistence, and primary non-adherence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Primary non-adherence: estimated at 11.6%. These only incurred the cost of screening</td>
</tr>
<tr>
<td>Study</td>
<td>Question</td>
<td>Summary</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Jansen (2008)(^{186})</td>
<td>Persistence:</td>
<td>42.5% of those who initiated treatment discontinued within 6 months. These were given no treatment benefit but incurred 3 months treatment cost. Another 18.1%, 13.9% and 7.2% were dropped off therapy at 1 year, 2 years and 3 years, respectively. Patients discontinuing therapy received no further treatment.</td>
</tr>
<tr>
<td>Compliance: patients were considered compliant if their MPR was ≥ 80%. The probabilities of being less than 80% compliant were 23.9%, 4% and 1.2% in the first, second and following years of therapy, respectively. From the Belgian study, poor compliance was associated with an increased fracture rate of 35% for hip fractures. From an alternative source, it was assumed that for non-hip fractures the increase in fracture rate would only be 17%. The relative risks of fracture from the meta-analysis were associated with a population with a MPR of ≥ 80%, so if alendronate was assumed to reduce the risk of hip fracture by 38% then non-compliant women would receive a reduction in hip fractures of 16% (0.62 × 1.35 = 0.837). For poorly compliant women, the drug cost was restricted to 80% of full price.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis: because the adherence rates varied by region, additional analyses were conducted assuming that adherence rates were 20% and 40% higher than in the real-world scenario.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>Country?</td>
<td>UK and the Netherlands</td>
<td></td>
</tr>
<tr>
<td>Type of model?</td>
<td>This was an individual patient simulation model, a replicate of the Markov health-state transition model developed by Kanis et al. (2002)(^{187}) and adapted by Stevenson et al. (2005)(^{188}).</td>
<td></td>
</tr>
<tr>
<td>Study objective?</td>
<td>To evaluate the cost-effectiveness of a fixed dose combination of alendronate and cholecalciferol versus no treatment, alendronate treatment with dietary vitamin D supplements and ibandronate in the treatment of osteoporosis.</td>
<td></td>
</tr>
<tr>
<td>Adherence definition?</td>
<td>Adherence (treatment compliance rate) only Persistence regarding bisphosphonates was not taken into account.</td>
<td></td>
</tr>
<tr>
<td>How was adherence modelled?</td>
<td>At baseline, it seems that full compliance was assumed in each arm. In additional analyses, in order to account for adherence, effectiveness for Vitamin D was increased by multiplying the relative risk by 0.9, and non-compliance was assumed to be 30%. This was not satisfactorily explained.</td>
<td></td>
</tr>
<tr>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>Kanis (2002)(^{187})</td>
<td>Country?</td>
<td>UK</td>
</tr>
<tr>
<td>Type of model?</td>
<td>Markov microsimulation</td>
<td></td>
</tr>
<tr>
<td>Study objective?</td>
<td>The cost-effectiveness of various agents for the treatment of established osteoporosis is modelled.</td>
<td></td>
</tr>
<tr>
<td>Adherence definition?</td>
<td>Adherence is not mentioned Compliance is distinguished from continuance.</td>
<td></td>
</tr>
<tr>
<td>How was adherence modelled?</td>
<td>Continuance is the duration of taking the treatment Compliance was the proportion of the medication being taken by the patient Non-compliant patients were assumed to incur 3 months’ worth of medication costs and receive no treatment benefit. In the base case, 100% compliance was assumed and this was varied in sensitivity analysis.</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Question</td>
<td>Summary</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Majumdar (2007)</td>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
<tr>
<td></td>
<td>Country?</td>
<td>Canada</td>
</tr>
<tr>
<td></td>
<td>Type of model?</td>
<td>Decision tree (1 year) and Markov model (post 1 year to lifetime)</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>To examine longer-term outcomes, reproducibility and cost-effectiveness of a multifaceted intervention to improve the quality of osteoporosis care after fracture of the wrist, which involved encouraging patients to come forward and be treated through physician reminders, local opinion leader endorsed treatment guidelines, and patient guidelines</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>Only persistence was incorporated in the model. A patient was considered persistent if they were filling in their prescriptions</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>In the model, based on a clinical study, 1-year persistence of osteoporosis treatment was 80% and this was assumed to continue for the next 4 years. It was also assumed that the 20% of patients who discontinued treatment did so in the first year and that they received no fracture reduction benefits whatsoever</td>
</tr>
<tr>
<td></td>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Patrick (2011)</td>
<td>Country?</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>Type of model?</td>
<td>A microsimulation, state transition model</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>The objective was to model different medication adherence patterns among women initiating bisphosphonate treatment and to estimate the cost-effectiveness of a hypothetical intervention to improve adherence</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>There was no definition of adherence; patients were considered to be on or off treatment. There was no switching treatments</td>
</tr>
<tr>
<td></td>
<td>How was adherence modelled?</td>
<td>The probabilities of treatment discontinuation (having a 30-day gap during which no treatment was available) and reinitiation were included in the model</td>
</tr>
<tr>
<td></td>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
</tr>
<tr>
<td></td>
<td>Type of model?</td>
<td>An individual state transition model</td>
</tr>
<tr>
<td></td>
<td>Study objective?</td>
<td>To develop a modelling framework that incorporates variables associated with adherence, and to identify important drivers of cost-effectiveness to inform future studies of adherence in osteoporosis</td>
</tr>
<tr>
<td></td>
<td>Adherence definition?</td>
<td>Adherence: a general term encompassing all aspects of persistence, compliance, and primary non-adherence. Persistence: the duration of therapy. The number of days until discontinuation of the proportion of the cohort still on medication after a given time. Compliance: proximity to the recommendations of the optimal treatment. This includes how long a drug is taken and can be simplified as the number of doses taken, divided by the number of prescribed doses during a defined period. (The term compliance also includes other aspects such as if a drug should be taken with or without food, the time of day it should be taken, whether or not doses are taken to compensate for forgotten doses, drug vacations, pill dumping, etc.)</td>
</tr>
</tbody>
</table>
### Study Question Summary

<table>
<thead>
<tr>
<th>Study</th>
<th>Question</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary non-adherence: if patients are prescribed a drug and never fill the prescription they are termed a primary non-adherent</td>
<td>Patients on treatment were classified as fully adherent or partially adherent. A fully adherent patient would receive treatment for the prescription duration. These patients received the expected treatment benefit of a fully compliant patient. A partially adherent patient was at risk of dropping out of treatment and had only a fraction of the benefit that a fully compliant patient would have.</td>
<td>The risk of dropping out of treatment would apply within the first 3 years; thereafter, persistence would be stable. Non-parametric dropout rates were obtained from a US database. A sensitivity analysis of different rates for different parts of the world was conducted. If a patient dropped out within 6 months, they received no treatment benefit but the cost of physician visits, BMD measurements and 3 months of drug costs were incurred.</td>
</tr>
<tr>
<td>How was adherence modelled?</td>
<td>It is stated that in most large clinical trials, ≥ 80% of patients are persistent until end of trial, but no adjustments were made here for the adherent group because of its conceptual context. Using MPR as a measure of compliance, studies have estimated differences in fracture rates between compliant and non-compliant patients to range between 16% and 44%. However, non-compliant patients have higher comorbidities, are more frail, and have higher medical expenditure than compliant patients and fracture rates are higher in non-compliant patients taking placebo. As these estimates are seldom controlled for clinical risk factors, a base-case fraction of the benefit of 80% was assumed and sensitivity analysis was done between 0% and 100%</td>
<td>Primary adherence was set to 4%. They were assumed to incur the cost of one physician visit and one BMD measurement.</td>
</tr>
<tr>
<td>How was treatment management modelled?</td>
<td>Not applicable</td>
<td></td>
</tr>
</tbody>
</table>

BM, bone turnover marker; BP, biophosphonate; MPR, medical possession ratio.
Appendix 6 Final protocol

1. Title of the project:
Bone turnover markers for monitoring the response to osteoporosis treatment: the secondary prevention of fractures, and primary prevention of fractures in high risk groups.

2. Name of TAR team and ‘lead’
Centre for Reviews and Dissemination/Centre for Health Economics Technology Assessment Group, University of York.

Summary
Bone turnover markers may be useful for identifying patients with osteoporosis who are not responding to treatment, which in turn will allow changes in management or treatment strategies to be implemented in a timely manner to ensure maximum benefit to the patient. An evidence synthesis using systematic review methodology will be used to investigate potential uses of bone turnover markers, and a decision analytical model developed if sufficient evidence is found to establish clinical effectiveness.

5. Decision problem
The review of the clinical evidence will focus on three key clinical areas:

- Clinical effectiveness: how does bone marker monitoring impact on the decision making process and patient outcomes?
- Test accuracy: how well do the results of the biomarker tests correlate with changes in bone density, architecture and incidence of fracture?
- Test reliability and reproducibility: how much do the results of tests vary within and between patients?

If clinical effectiveness can be established, a decision modelling will be developed and a expected value of perfect information (EVPI) analysis undertaken. Any EVPI analysis is dependent on the ability to undertake decision modelling. The decision model will focus on the effect of bone marker testing on patient management decisions, and will address the question: ‘Which monitoring regimen is the most cost-effective in informing treatment decision.’ The treatments being considered are bisphosphonates (oral and intravenous), raloxifene, strontium ranelate, teriparatide, denosumab and no treatment.

6. Objectives
The primary aims of the systematic review are to determine the clinical effectiveness, test accuracy, test reliability and test reproducibility, of bone turnover markers in people being treated with any bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for osteoporosis. If possible, a decision model will be developed to determine the cost-effectiveness of bone turnover markers for monitoring treatment response and making changes in patient management. If a decision model is produced, EVPI analyses will be used to determine the need for further research, identify the research questions critical to decision making, and help inform the design of future studies and to consider implementation issues.
7. Methods of synthesising evidence of clinical effectiveness

The review will be conducted systematically following the general principles recommended in CRD guidance for undertaking reviews in health care and the PRISMA statement.

Search strategy
The following databases will be searched to identify primary studies, relevant reviews and economic studies:

- CINAHL
- Cochrane Library (including the Cochrane Database of Systematic Reviews, Database of Reviews of Abstracts of Effects (DARE), Health Technology Assessment (HTA) Database, NHS Economic Evaluation Database (NHS EED), and the Cochrane Central Register of Controlled Trials)
- EconLit
- EMBASE
- MEDLINE
- Science Citation Index

The following sources will be searched to identify grey literature and ongoing research:

- Clinical Trials.gov
- Conference Proceedings Citation Index - Science
- Controlled Clinical Trials.com

A draft search strategy for use with MEDLINE is provided in Appendix 1. No language or date restrictions will be applied during the search. Additional searches will be conducted as required.

Inclusion and exclusion criteria

Population
Studies eligible for inclusion will be those in adults (> 18 years of age) either:

- Receiving any bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for the secondary prevention of osteoporotic fractures, regardless of the baseline pathology, or
- In any high-risk group being treated with any bisphosphonate, raloxifene, strontium ranelate, denosumab or teriparatide for the primary prevention of osteoporotic fractures.

Interventions
P1NP (serum), CTX (urinary and serum), NTX (urinary and serum), and BAP (serum).

Study designs
Effectiveness: RCTs where patients are randomised to a standard monitoring regimen (with or without DXA), or to standard monitoring regimen with additional monitoring with a bone turnover marker. Studies reporting the impact of bone marker test results on the decision making process for management of osteoporosis, that also report the subsequent rate of fracture in the population being assessed, will also be sought (‘Decision studies’).

Test accuracy: Studies comparing the results of bone marker tests to the results of bone biopsy or a composite reference standard of DXA and subsequent fracture outcome will be included. Given the nature of the review question, we believe it is unlikely that such studies will be available. So in addition we will include prospective studies that measure the association between bone turnover and bone density and/or fracture rates, and that report a correlation coefficient for this association. Prospective studies that evaluate changes in bone biomarkers in patients receiving one of the specified osteoporosis treatments, that
provide sufficient data to produce a measure of the risk of fracture, or that report the results of multivariate regression analyses in which a biomarker of interest is an independent variable, will also be eligible for inclusion.

Reliability and reproducibility: Prospective controlled studies of serial bone marker measurements that report a measure of within and/or between patient variability, will be included.

Studies assessing the effectiveness of treatments for osteoporosis using changes in bone turnover biomarkers solely as an outcome will be excluded. Prognostic studies using biomarkers to identify patients at risk of osteoporosis and fracture at baseline, prior to commencing treatment, will also be excluded.

Outcomes

**Effectiveness**

RCTs and decision studies reporting either change in patient management strategies, the incidence of fracture and/or treatment adherence rates.

**Test accuracy**

Studies will have to report either:

- Estimates of diagnostic accuracy or sufficient data for these to be calculated
- A correlation coefficient, or sufficient data for this to be calculated, for the association between a bone turnover marker and bone density and/or the incidence of fracture
- The risk/incidence of fracture associated with the bone marker test results
- At least a p-value for a bone marker of interest that is used as an independent variable in a multivariate regression.

**Reliability and reproducibility**

Studies reporting a measure for intra- and/or inter-patient variability in bone marker test results.

**Data extraction strategy**

Data extraction will be conducted by one reviewer using a standardised data extraction form and checked by a second reviewer. Discrepancies will be resolved by discussion, with involvement of a third reviewer when necessary. If time constraints allow, attempts will be made to contact authors for missing data. Data from multiple publications of the same study will be extracted and reported as a single study. Where applicable and available, extraction will include data on: study details (e.g. study identifier/EndNote ID, author, year, country, setting, number of participants, and duration of follow up), patient characteristics (e.g. age, gender, ethnicity, duration of osteoporosis, risk group, concomitant renal/liver disease; baseline P1NP, CTX and/or NTX levels), details of intervention (serum or urine, sample collection details; pre-sampling preparations/restrictions; sample storage details; assay used; adjustments for creatinine excretion; delay between sample collection and assay; single/serial measures; thresholds/cut-offs/reference values), study quality, and reported outcomes as specified above.

**Quality assessment strategy**

The quality of the individual studies will be assessed by one reviewer, and independently checked by a second reviewer. Any disagreements will be resolved by consensus and if necessary a third reviewer will be consulted. The quality of included studies will be assessed using standard checklists suitable for the study design, and adapted as necessary to incorporate topic-specific quality issues.
Methods of analysis/synthesis

Key study characteristics, patient outcomes and study quality will be summarised in a narrative and tables. Where appropriate, meta-analysis suitable to the data extracted will be employed to estimate a summary measure of effect based on intention to treat analyses. Potential sources of heterogeneity will explored:

- Subgroups of potential interest will be investigated if sufficient data are available, for example, post-menopausal women (overall and for specific age ranges if data are available), elderly, skeletal site (hip, spine, wrist), and glucocorticoid-induced osteoporosis
- Sensitivity analyses will be conducted, where appropriate, to investigate potential sources of heterogeneity such as study quality, and differences in sample acquisition, storage and assay methods.

8. Methods of synthesising evidence of cost-effectiveness

- Identifying and systematically reviewing published cost-effectiveness studies

Systematic searches will be undertaken to identify existing published studies reporting the cost-effectiveness of bone-turnover markers for monitoring the response to osteoporosis treatment. The following databases will be searched: MEDLINE, EMBASE, CENTRAL and EconLit. In addition, searches of NHS EED and HEED will be carried out, along with a search of the Economics Working Papers archive (IDEAS).

Only full economic evaluations that compare two or more options, that meet the inclusion criteria for the clinical review and consider both costs and consequences (including cost-effectiveness, cost-utility and cost-benefit analyses), will be included in the review of economic literature.

The quality of the cost-effectiveness studies will be assessed according to a checklist updated from that developed by Drummond et al. (2005) and Philips et al. (2002). This checklist will reflect the criteria for economic evaluation detailed in the methodological guidance developed by the National Institute for Health and Clinical Excellence. This information will be tabulated and summarised within the text of the report. In particular, information will be extracted on the comparators, study population, main analytic approaches (e.g. patient-level analysis/decision-analytic modelling), primary outcome specified for the economic analysis, details of adjustment for quality-of life, direct costs (medical and non-medical) and productivity costs, estimates of incremental cost-effectiveness and approaches to quantifying decision uncertainty (e.g. deterministic/probabilistic sensitivity analysis).

The review will examine the full economic evaluations that meet the inclusion criteria in detail, with the aim of identifying important structural assumptions, highlighting key areas of uncertainty and outlining the potential issues of generalising from the results of existing economic evaluations.

- Development of a new decision-analytic model

If relevant effectiveness evidence can be identified (this may be in the form of an effect measure from an RCT or an appropriate predictive value from a test accuracy study), a decision-analytic model will be developed to determine the cost-effectiveness of bone-turnover markers for monitoring treatment response and informing changes in patient management. One possibility is to use an existing peer-reviewed decision model developed by ScHARR (University of Sheffield) to estimate the cost-effectiveness of osteoporosis interventions, using the most recent work undertaken. The model developer has agreed to provide access to this model for the purposes of this project (Personal communication: Dr Matt Stevenson). However, potential issues of academic in confidence data will need to be clarified before determining the final version of the model which will be used. If monitoring clinical effectiveness data and adherence data are identified then the Sheffield meta-model could be utilised. The Sheffield meta-model is a simpler summary model of the original individual patient simulation (IPS) model. Cost data in the model will be updated using the most contemporary estimates from national databases (e.g. reference costs), and a literature review will be
conducted to identify any relevant utility estimates in addition to those used in the existing model. Discounting will be undertaken at an annual rate of 3.5% on costs and benefits.

If test accuracy data is available and it is possible to utilise these data in the original IPS model then this will also be considered. Additional searching will be undertaken, if required, to identify relevant model structures from published cost-effectiveness analyses. These will be used to help inform this adaption of the IPS model. Further, if the use of the Sheffield model is not an option the published models identified will be utilised in the development of a new decision model.

The presence of any data gaps (e.g. resource use data) that may need to be filled during the development of the model will be identified from the literature identified during the systematic review process and additional searches if required. The primary outcome of the model will be the cost-utility of different monitoring strategies. The number of fractures prevented will also be reported. Cost-effectiveness will be established by estimating incremental cost-effectiveness ratios. The number of fractures prevented will also be reported. The risk-benefit uncertainties such as the clinical effect, adverse event and net-benefit uncertainties, and the model assumptions will be presented clearly.

To consider future research priorities in the NHS, the model will also be used to undertake analyses of the EVPI. Depending on whether a model is built on the fracture risk clinical effectiveness of monitoring strategies or test accuracy, EVPI analyses will be conducted for the relevant data in the model. EVPI represents the expected costs of decision uncertainty since perfect information would eliminate the possibility of making the wrong decision. Hence, EVPI for the overall decision problem represents the value of eliminating all uncertainty and EVPI for key parameters (termed partial EVPI) represents the value of eliminating uncertainties in particular subsets of parameters. Separate analyses will be undertaken to reflect the variability considered in the decision model itself if the model allows. Per patient EVPI estimates will be scaled up to reflect the relevant UK population size and will adopt an appropriate time-horizon. EVPI also represents the maximum amount that a decision-maker should be willing to pay for additional evidence to inform this decision in the future. EVPI provides an upper bound on the value of additional research. The objective of this analysis (termed partial EVPI) is to identify the model parameters where it would be most worthwhile obtaining more precise estimates. The results from the clinical effectiveness review and the EVPI results will be used to identify future research recommendations.

9.7 TAR Centre

The Technology Assessment Review team at the University of York is drawn from two specialist centres: the Centre for Reviews and Dissemination (CRD) and the Centre for Health Economics (CHE). This Technology Assessment will be conducted by CRD.

CRD undertakes reviews of research about the effects of interventions used in health and social care (www.york.ac.uk/inst/crd). The centre maintains various databases, provides an enquiry service and disseminates results of research to NHS decision makers.

Recent TARs undertaken by CRD/CHE at York relate to the identification of the seizure focus in patients with refractory epilepsy being considered for surgery, aldosterone treatment for post-MI heart failure, treatments for bipolar disorder, sugammadex for the reversal of muscle relaxation in general anaesthesia and photodynamic therapy in the treatment of specified cancer sites.
10 Expertise in the TAR team and team contributions

Jane Burch, Research Fellow (jane.burch@york.ac.uk). Eight years’ experience in systematic reviews and systematic review methodology. Has worked on systematic reviews for NICE, the HTA programme and the NHS Cancer Screening Programmes. Will be responsible for all aspects of the clinical effectiveness review and co-ordinating the production of the final report.

Stephen Rice, Research Fellow in Health Economics (stephen.rice@york.ac.uk). Over seven years’ experience in economic evaluation and evidence synthesis. Will be responsible for the cost-effectiveness review, development of any cost-effectiveness model, and writing the economic sections of the report.

Aileen Neilson, Research Fellow in Health Economics (aileen.neilson@york.ac.uk). Involved with various health outcomes research and economic evaluation studies within the National Health Service setting in the UK, and against a broader European context. Will assist with the cost-effectiveness review, development of any cost-effectiveness model, and writing the report.

Huiqin Yang, Research Fellow (huiqin.yang@york.ac.uk). Six years’ experience in health services research. Has worked on systematic reviews for NICE and the HTA programme. Will assist with all aspects of the clinical effectiveness review and the writing of the final report.

Lisa Stirk, Information Officer (lisa.stirk@york.ac.uk). Over twelve years’ experience in literature searching for systematic reviews. Has worked on systematic reviews for NICE, the HTA programme and the British Thoracic Society. Will be responsible for devising the search strategy, carrying out the literature searches and maintaining the literature database.

Professor Roger Francis, Emeritus Professor of Geriatric Medicine, Institute for Ageing and Health, Newcastle University (r.m.francis@newcastle.ac.uk) and formerly Consultant Physician, Bone Clinic at Freeman Hospital, Newcastle upon Tyne. Involved in clinical research related to osteoporosis for 30 years and will provide clinical advice throughout the project commenting on the protocol, results and report.

Dr Paul Holloway (paul.holloway@imperial.ac.uk). Clinical and academic interest in metabolic bone disease since training as senior registrar and clinical lecturer in Oxford in 1980s. Has run a metabolic bone clinic at St Mary’s since 2004 and is acting director of the St Mary’s SAS for bone markers. Will provide advice and comments on the protocol and report.

Dr Peter Selby, Consultant Physician, Manchester Royal Infirmary, Honorary Senior Lecturer, University of Manchester (peter.selby@manchester.ac.uk). Involved in management of patient with osteoporosis and clinical research in bone disease for over 25 years and will provide clinical advice throughout the project, commenting on the protocol, results and report.

Dawn Craig, Research Fellow (dawn.craig@york.ac.uk). Over eight years’ experience in economic evaluation and health technology assessment in a wide variety of areas. Contributed to the drafting of the protocol and will provide input at all stages of the project and comment on draft/final report. Has overall responsibility for the management of both the clinical and economic components of the project.

Advisory group

Professor John Kanis (w.j.pontefract@sheffield.ac.uk). An expert on metabolic bone diseases and director of the WHO Collaborating Centre for Metabolic Bone Diseases at Sheffield. He has a long experience in Health Technology Assessment, guideline development and WHO Scientific Study Group reports. Will provide advice and comments on the report.