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 Selby[4] and Dawn Craig[1]*

[1]Centre for Reviews and Dissemination, York, UK
[2]Institute for Ageing and Health, Newcastle University, Newcastle upon Tyne, UK
[3]Imperial College Healthcare Trust, St Mary's Hospital, London, UK
[4]Department 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

Scientific summary

Bone turnover markers for osteoporosis treatment

Health Technology Assessment 2014; Vol. 18: No. 11
DOI: 10.3310/hta18110

NIHR Journals Library www.journalslibrary.nihr.ac.uk

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

© 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.

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 maker 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.

© 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.

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.

© 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.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.804

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

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