A systematic review and economic evaluation of bisphosphonates for the prevention of fragility fractures

Sarah Davis,[1]* Marrissa Martyn-St James,[1] Jean Sanderson,[1] John Stevens,[1] Edward Goka,[1] Andrew Rawdin,[1] Susi Sadler,[1] Ruth Wong,[1] Fiona Campbell,[1] Matt Stevenson,[1] Mark Strong,[1] Peter Selby[2] and Neil Gittoes[3]

[1]Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK
[2]Department of Medicine, University of Manchester, Manchester Royal Infirmary, Manchester, UK
[3]Centre for Endocrinology, Diabetes and Metabolism, Birmingham Health Partners, University Hospitals Birmingham, Birmingham, UK

*Corresponding author

Declared competing interests of authors: Peter Selby reports personal fees from Internis Pharma and non-financial support from AMGen, outside the submitted work. Neil Gittoes reports personal fees from the advisory board at Eli Lilly and Company, personal fees from the advisory board at AMGen, personal fees from speaker fees at AMGen, personal fees from speaker fees at GlaxoSmithKline, personal fees from the advisory board at ProStrakan, personal fees from advisory board at Shire, personal fees from the advisory board at Internis Pharma, personal fees from the advisory board at Consilient Health and personal fees from the advisory board at NPS Pharmaceuticals, outside the submitted work.

Published October 2016
DOI: 10.3310/hta20780

Scientific summary

Bisphosphonates for the prevention of fragility fractures

Health Technology Assessment 2016; Vol. 20: No. 78
DOI: 10.3310/hta20780

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

Scientific summary

Background

Osteoporosis is a disease characterised by low bone mass and structural deterioration of bone tissue, with a consequent increase in susceptibility to fragility fracture. Fragility fractures are fractures that result from mechanical forces that would not ordinarily result in fracture, known as low-level (or 'low-energy') trauma. The World Health Organization (WHO) has quantified this as forces equivalent to a fall from a standing height or less. Although osteoporosis is an important predictor of the risk of fragility fracture, 70% of fragility fractures in postmenopausal women occur in those who do not meet the criteria for osteoporosis. Every year 300,000 people in the UK suffer a fragility fracture, of which over 70,000 are hip fractures.

Objectives

To evaluate the clinical effectiveness, safety and cost-effectiveness of bisphosphonates for the prevention of fragility fractures.

Methods

A systematic review of the literature including network meta-analyses (NMA) was conducted in order to evaluate the clinical effectiveness and safety of oral [alendronic acid (Fosamax® and Fosamax® Once Weekly, Merck Sharp & Dohme Ltd), ibandronic acid (Bonviva®, Roche Products Ltd) and risedronic acid (Actonel® and Actonel Once a Week®, Warner Chilcott UK Ltd)] and intravenous (i.v.) [ibandronic acid and zoledronic acid (Aclasta®, Novartis Pharmaceuticals UK Ltd)] bisphosphonates in the prevention of fragility fractures. For the clinical effectiveness review, six electronic databases and two trial registries were searched: MEDLINE, EMBASE, The Cochrane Library, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science and BIOSIS Previews, Clinicaltrials.gov and WHO International Clinical Trials Registry Platform. Searches were limited by date from 2008 until September 2014. A review of the existing cost-effectiveness literature was undertaken. In the cost-effectiveness review (economic evaluation and quality-of-life studies), seven electronic databases were searched from 2006 until September 2014: MEDLINE, EMBASE, The Cochrane Library, CINAHL, EconLit, Web of Science and BIOSIS Previews. Additional searches were carried out in October 2014–January 2015 in MEDLINE and EMBASE for adverse events, compliance and EuroQol five dimensions questionnaire to inform the model parameters. A de novo health economic model was constructed using discrete event simulation in order to evaluate the cost-effectiveness of the interventions under assessment.

Results

Number and quality of studies

A total of 46 randomised controlled trials (RCTs) were identified for the clinical effectiveness systematic review. Alendronic acid was evaluated against placebo in 17 RCTs. Daily oral ibandronic acid (unlicensed dose) was evaluated against placebo in three RCTs and against i.v. administration in one RCT. Daily administration of oral ibandronic acid was evaluated against monthly oral administration in one RCT. Risedronic acid was evaluated against placebo in 12 RCTs and zoledronic acid was evaluated against placebo in four RCTs. One RCT evaluated alendronic acid compared with monthly oral ibandronic acid, five RCTs compared alendronic acid with risedronic acid, one RCT compared zoledronic acid with alendronic acid and one RCT compared zoledronic acid with risedronic acid.

The risk of bias associated with the included RCTs was assessed using the Cochrane risk-of-bias instrument. Attrition ≥ 10% across treatment groups was evident for 29 (63%) of the included RCTs. Five trials were reported as either open label or single blind and were considered at high risk of performance bias. Blinded outcome assessment was reported by only 13 (28%) trials.

Summary of benefits and risks

The outcome measures prespecified in the final National Institute for Health and Care Excellence scope were addressed by the included trial evidence to varying degrees. Femoral neck bone mineral density (BMD) was the most widely reported outcome; fracture was the second most widely reported outcome. Adverse events (AEs) were reported by the majority of included trials. Across the included trials there was limited reporting on the outcomes of compliance (adherence and persistence), hospitalisation and service use, and quality of life.

A total of 27 RCTs provided suitable fracture data for inclusion in the NMA and a total of 35 RCTs provided suitable femoral neck BMD data for inclusion in the BMD NMA. Based on the NMA, all treatments were associated with beneficial effects relative to placebo, with hazard ratios varying from 0.41 to 0.92 depending on treatment and fracture site. For vertebral fractures and percentage change in BMD, the treatment effects were also statistically significant at a conventional 5% significance level for all treatments. Pairwise comparisons between bisphosphonates indicated that no bisphosphonate was statistically significantly more effective than any other bisphosphonate for fracture outcomes. For vertebral fractures and BMD, the greatest effect was for zoledronic acid, although, in general, the ranking of treatments varied for the different outcomes, with the treatments providing broadly similar effects.

Assessment of vertebral fractures within the trials was based on both clinical and morphometric fractures. Ideally, the effect of assessment method would have been assessed using metaregression, but there were insufficient data to facilitate this. Consideration of the trials reporting clinical fractures did not provide any evidence to suggest significantly different treatment effects according to assessment method.

Pooled RCT data for each bisphosphonate indicated no statistically significant differences in the incidence of upper gastrointestinal (GI) events, no evidence of significant differences in mortality and no significant differences in participants withdrawing because of AEs. Evidence from single RCTs indicated that the risk of upper GI events was significantly higher in men receiving risedronic acid than in those receiving placebo, that men and women receiving placebo were significantly more likely to die following hip fracture than those receiving zoledronic acid, and that the proportion of men withdrawing because of AEs was significantly higher among those receiving alendronic acid than among those receiving placebo.

Pooled RCT data indicated evidence of influenza-like symptoms associated with zoledronic acid. Single RCT evidence indicated no statistically significant difference in the incidence of atrial fibrillation, bone pain or stroke. Single RCT evidence indicated a statistically significant risk of eye inflammation in the first 3 days following administration of zoledronic acid. All RCTs evaluating zoledronic acid reported no cases of spontaneous osteonecrosis of the jaw.

Adverse events of hypocalcaemia and atypical femoral fracture were not reported outcomes in any RCT of any bisphosphonate.

Summary of cost-effectiveness evidence

The de novo economic model estimates that a strategy of no treatment is predicted to have the greatest net benefit for patients, with an absolute risk < 1.5% when using QFracture® (QFracture-2012 open source revision 38, Clinrisk Ltd, Leeds, UK) to estimate absolute risk and valuing a quality-adjusted life-year (QALY) at £20,000. Alendronic acid is predicted to have the maximum incremental net benefit (INB) from 1.5% to 7.2% and risedronic acid is predicted to have the maximum INB from 7.2% upwards. However, the absolute costs and QALY gains are small in patients with low absolute risk and the probabilistic sensitivity analysis (PSA) suggested that there is considerable uncertainty regarding whether or not no treatment is the optimal strategy until the QFracture score is around 5.5% (the mean absolute risk for the eighth risk category for QFracture).

© Queen's Printer and Controller of HMSO 2016. This work was produced by Davis *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.

The mean INBs for oral bisphosphonate treatment (alendronic acid, risedronic acid and ibandronic acid) compared with no treatment were positive across all FRAX® (web version 3.9; University of Sheffield, Sheffield, UK) risk categories. An exact threshold for the absolute risk at which the INB became positive was therefore not available but the minimum FRAX score in the modelled population was 1.2% and the lowest risk category (containing one-tenth of the modelled population) had a mean absolute risk of 3.1%. Oral ibandronic acid is predicted to have the highest INB compared with no treatment up to 8.6%, with alendronic acid having the highest INB from 8.6% to 38.5% and risedronic acid having the maximum INB above 38.5%. The PSA suggested that there was a low probability of the no-treatment strategy being optimal across all FRAX risk categories when valuing a QALY at £20,000. However, the PSA also demonstrated that there is considerable uncertainty regarding the optimal bisphosphonate treatment, with all of the oral bisphosphonates having reasonably similar probabilities of having maximum INB across most of the FRAX risk categories.

Intravenous bisphosphonates (ibandronic acid and zoledronic acid) were predicted to have lower INBs than oral bisphosphonates across all levels of absolute risk when estimated using either QFracture or FRAX. In the highest risk categories the incremental cost-effectiveness ratios for i.v. ibandronic acid and i.v. zoledronic acid compared with oral bisphosphonates were consistently > £50,000 per QALY, even though the base-case analysis assumed that patients treated with i.v. bisphosphonates persisted with treatment for longer than patients treated with oral bisphosphonates. Although the mean INB compared with no treatment for i.v. ibandronic acid did become positive at very high levels of absolute risk when using QFracture, the results when using FRAX went in the opposite direction. This may be as a result of the small number of patients and parameter samples informing the estimates at high levels of absolute risk, which makes these estimates more uncertain.

The results appeared to be broadly similar across the majority of the structural sensitivity analyses that examined the application of alternative data or assumptions. The results were more favourable to treatment when assuming that participants persisted with treatment for the full intended treatment duration (3 years for zoledronic acid and 5 years for all other bisphosphonates) or when assuming no AEs. The sensitivity analysis examining an AE rate of 30% in the month following initiation of oral bisphosphonate therapy showed that the cost-effectiveness of oral bisphosphonates is very sensitive to the rate of AEs experienced, with the INBs for oral bisphosphonates versus no treatment falling below zero (when valuing a QALY at £20,000) for all 10 QFracture risk categories and for all but the highest FRAX risk category.

Two structural sensitivity analyses that varied the way in which the fracture risk was estimated showed results that were broadly similar for QFracture but slightly less favourable for FRAX. In these sensitivity analyses, the cost-effectiveness estimates from the QFracture and FRAX model were closer together for patients with similar mean absolute risk than in the base case.

Discussion

Strengths, limitations of the analyses and uncertainties

The clinical effectiveness systematic review was based on rigorous methods, with comprehensive searches for evidence (up to September 2014), a good level of consistency between reviewers in study selection and double-checking of data extraction. A formal assessment of methodological quality of included trials was undertaken. Attrition ≥ 10% across treatment groups was evident for 63% of the included RCTs.

Not all of the included studies provided data suitable for inclusion in the NMA. For fracture there was variability across the included trials in the skeletal fracture site evaluated, the most frequently evaluated being vertebral fracture. Femoral neck BMD summary statistics were not provided by all trials but were extracted from graphical representations where possible. NMAs were performed to permit a coherent comparison of the efficacy of interventions in terms of fracture and femoral neck BMD.

Adverse event data were widely reported in the included RCTs and supplemented by review evidence of observational data. Evidence for compliance and persistence was mainly limited to review evidence of observational data.

The Assessment Group's economic analysis has a number of strengths.

- The patient-level simulation approach used in the economic model allowed for the distribution of patient characteristics to differ across the risk categories providing estimates of cost-effectiveness that have taken into account the differing consequences of fracture in patients with different characteristics.
- The economic modelling approach used allowed intervention thresholds to be linked to absolute risk measured using the two risk assessment tools recommended in clinical guideline 146 (CG146) (National Institute for Health and Care Excellence. *Osteoporosis: Assessing the Risk of Fragility Fracture*. Manchester: National Institute for Health and Care Excellence; 2014) as specified in the scope.
- Non-parametric regression was used to estimate the relationship between INB and absolute risk when averaging over both parameter uncertainty and the stochastic uncertainty associated with patient-level simulations.
- The economic model was underpinned by a NMA across all drug options, which provided a consistent framework for synthesising relevant efficacy data within a single network of evidence.

The Assessment Group's economic model is also subject to a number of limitations.

- In order to provide a single intervention threshold for each treatment that could be applied across the whole population, we had to assume that all of the bisphosphonate treatment strategies were viable treatment options in all patients eligible for risk assessment within CG146. This would not be true if the licensed indications for each intervention were to be strictly applied. Furthermore, the studies included in the NMA that informed the economic evaluation are not strictly exchangeable because not all interventions are licensed in all patient populations.
- The cost-effectiveness of treatment in the lowest-risk categories was particularly sensitive to the assumptions regarding the adverse effects of treatment because of the low absolute QALY gains and cost savings attributable to prevented fractures.
- The results of structural sensitivity analyses suggest that the model using FRAX to estimate absolute risk may have overestimated the INB of treatment because of the assumption that the proportion of fractures occurring at the hip is similar for QFracture and FRAX.

Key uncertainties in this assessment include:

- There was no evidence of differential treatment effects with respect to sex and age. However, there was some heterogeneity in treatment effects between studies, suggesting differential treatment effects according to study characteristics and the effect of treatment on femoral neck BMD depended on the baseline response.
- It is uncertain whether or not the cost-effectiveness of bisphosphonate treatment at a particular level of absolute fracture risk would be similar for patients who have been assessed using the FRAX algorithm for patients with known BMD.
- The incidence of upper GI AEs following initiation of oral bisphosphonate treatment is uncertain as the findings differ between the RCT evidence and the observational evidence from prescription event monitoring studies.

Conclusions

All treatments were associated with beneficial effects relative to placebo. Pairwise comparisons between treatments indicated that no active treatment was significantly more effective than other active treatments in reducing fracture outcomes. Bisphosphonates are generally well tolerated in patients enrolled in clinical

© Queen's Printer and Controller of HMSO 2016. This work was produced by Davis *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.

trials but may be less well tolerated in clinical practice. Influenza-like symptoms are associated with treatment with zoledronic acid, although clinical advice was that these symptoms are generally limited to the first dose and usually last only a few days.

The de novo economic model estimates that, when using QFracture to estimate absolute risk, a strategy of no treatment is predicted to have the greatest net benefit, when valuing a QALY at £20,000, in the lowest-risk patients (QFracture absolute risk < 1.5%), with oral bisphosphonates having the greatest INB at higher levels of absolute risk. However, the absolute costs and QALY gains are small in patients with low absolute risk and the PSA suggested that there is considerable uncertainty regarding whether or not no treatment is the optimal strategy until the QFracture score is around 5.5% (the mean absolute risk for the eighth risk category for QFracture). Therefore, it is possible that patients and clinicians may not consider treatment worthwhile in the lowest-risk patients even though it may be cost-effective.

The mean INBs compared with no treatment (when valuing a QALY at £20,000) were positive for all oral bisphosphonate treatments across all FRAX risk categories. However, the results of two structural sensitivity analyses suggest that the base-case analysis may have overestimated the INBs of treatment in the model based on FRAX, owing to the assumption that the proportion of major osteoporotic fractures occurring at the hip is the same for FRAX and QFracture. Given this possible bias, and our belief that the results should be broadly similar across the two risk scores, it would be reasonable to assume that the absolute risk thresholds estimated in the QFracture model could be applied to patients whose score had been calculated using either QFracture or FRAX.

The de novo economic model suggests that the cost-effectiveness of i.v. bisphosphonates is less favourable than for oral bisphosphonates.

Further work is planned to extend the analysis to include non-bisphosphonate therapies.

Study registration

This study is registered as PROSPERO CRD42013006883.

Funding

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

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

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: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 13/04/001. The protocol was agreed in September 2014. The assessment report began editorial review in May 2015 and was accepted for publication in October 2015. 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 2016. This work was produced by Davis *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).

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME 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 Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

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

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, 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 Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

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

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, 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