UNIVERSITÀ DEGLI STUDI DI TORINO

Techniques for diagnosing osteoporosis: A systematic review of cost-effectiveness studies

Abstract

Objectives. The study question was whether dual-energy X-ray absorptiometry (DXA) alone is more cost-effective for identifying postmenopausal women with osteoporosis than a two-step procedure with quantitative ultrasound sonography (QUS) plus DXA. To answer this question, a systematic review was performed.

Methods. Electronic databases (PubMed, INAHTA, Health Evidence Network, NIHR, the Health Technology Assessment programme, the NHS Economic Evaluation Database, Research Papers in Economics, Web of Science, Scopus, and EconLit) were searched for cost-effectiveness publications. Two independent reviewers selected eligible publications based on the inclusion/exclusion criteria. Quality assessment of economic evaluations was undertaken using the Drummond checklist.

Results. Seven journal articles and four reports were reviewed. The cost per true positive case diagnosed by DXA was found to be higher than that for diagnosis by QUS+DXA in 2 articles. In one article it was found to be lower. In three studies the results were not conclusive. These articles were characterised by the differences in the types of devices, parameters and thresholds on the QUS and DXA tests and the unit costs of the DXA and QUS tests as well as by variability in the sensitivity and specificity of the techniques and the prevalence of osteoporosis.

Conclusions. The publications reviewed did not provide clear-cut evidence for drawing conclusions about which screening test may be more cost-effective for identifying postmenopausal women with osteoporosis.

Keywords: Comparative cost analysis; osteoporosis; diagnosis; economic evaluation

Introduction

Osteoporosis has become an increasingly recognised health concern by the medical community and the public. The hallmark of this skeletal disorder is diminished bone strength predisposing to a higher risk of fracture (1). Two types of osteoporosis are distinguished:

- primary osteoporosis, attributable to aging, menopause, and lifestyle-related factors, such as smoking, alcohol, diet and physical inactivity;
- secondary osteoporosis, caused by diseases and/or the use of drugs.

Primary osteoporosis affects millions of postmenopausal women and a growing number of men. Because of induced hormonal changes, it is more common among women after menopause. As such, it is perhaps the widest ranging social, physical and economic impact of oestrogen deficiency (2-4) and a leading risk factor for bone fractures in menopausal women (5). The incidence of osteoporotic fractures in Western countries is rising as the life expectancy lengthens. There is a clear relationship between bone mineral density (BMD) and fracture risk that facilitates the use of BMD as a predictive factor for the development of osteoporotic fractures. This approach, however, has two drawbacks: its predictive value is rather low in general (6), and its sensitivity further decreases with patients' increasing risk and age.

To achieve a higher sensitivity that is not affected by age, additional clinical risk factors independent of BMD, e.g., prevalent rheumatoid arthritis, smoking or excessive alcohol consumption, have been added to the evaluation.

Through this evaluation, an algorithm was developed that predicts the absolute 10-year fracture risk with a much higher predictive value than that from the evaluation of BMD or clinical risk factors alone (7,8).

The algorithm is known as FRAX® and is available free of charge at www.shef.ac.uk/FRAX®/. After the FRAX® algorithm is calibrated to local hip fracture and death rates, it is applicable to any geographic region (9).

According to a World Health Organisation (WHO) study group (10), the gold standard test for osteoporosis screening is the measurement of bone mineral density (BMD) by dual-energy X-ray absorptiometry (DXA or DEXA). Developed roughly 20 years ago (11), DXA is the method of choice for diagnosing osteoporosis and, consequently, fracture risk estimation.

Recently, there has been increased interest in the use of quantitative ultrasound sonography (QUS) (12,13) .

However, employing QUS to diagnose osteoporosis is somewhat problematic. The reason for this is that the WHO diagnostic classification applied to DXA T-scores cannot be used for QUS because QUS T-scores are not equivalent to T-scores derived by DXA (14); the explanation for this finding is that the two techniques measure distinct bone properties. Approaches to overcoming this dilemma require appropriate conversion equations and predefined, device-specific diagnostic thresholds; these, however, are still in development. Although osteoporosis screening by QUS is not recommended as a substitute for DXA, it may offer some potential advantages as a pre-test: QUS is easy to use, radiation free, and requires no special facilities for operation. For these reasons, QUS has been proposed as a pre-screening tool, with DXA offered only to those women identified by QUS as being at high risk for having osteoporosis (12,15,16). Consistent with the review of Schousboe (17) and the study by Nayak (18), consensus is lacking about the DXA test or the sequence QUS+DXA, the threshold for selecting individuals for treatment and the optimal age at which to initiate screening (17,18).

The policy question we posed for this study was whether DXA alone was more cost-effective for identifying postmenopausal women with osteoporosis than was a two-step procedure combining QUS with DXA. To answer this question, we performed a systematic review of the literature and evaluated the currently available evidence according to the PRISMA criteria (19).

This study is part of a research series in health technology assessment developed by the Department of Public Health and San Giovanni Battista University Hospital (Turin Italy). The research series focused on the performance and economic evaluations of different techniques (20,21).

Materials and Methods

Methods

Search strategy

In 2012, two researchers independently performed systematic searches of international databases to identify publications from PubMed, the International Network of Agencies for Health Technology Assessment [INAHTA], the Health Evidence Network [HEN], the National Institute for Health Research [NIHR] Health Technology Assessment programme, the National Health Service [NHS] Economic Evaluation Database, Research Papers in Economics [RePEc], the Web of Science, Scopus, and EconLit using MESH terms, text words and acronyms in multiple combinations.

All papers written in English, French and Italian, regardless of their dates of publication, were considered for our purposes.

Details of the search procedure are available in Supplementary Table 1.

Selection strategy and criteria

In the first stage, the researchers analysed the search results individually to find potentially eligible publications. The publications were sorted by title and abstracts; all irrelevant studies (lack of pertinence, identical publications found on more than one database) and reviews were excluded.

In the second phase, only the studies that met the following inclusion criteria were selected:

- the patients had to be postmenopausal women;
- the study had to compare QUS plus DXA with DXA alone. No exclusion criterion was applied to the publication types;
- the measurement of effectiveness had to be reported as the number of osteoporotic subjects accurately diagnosed, that is, the number of true positive cases;
- the publications had to contain sensitivity and specificity or allow for their calculation;

- the technique used had to be DXA at the femoral neck or lumbar spine or total hip (22) and calcaneal QUS;
- the economic evaluation of resources required to provide the alternative techniques (DXA and QUS test costs) had to be included.

The exclusion process was performed by two independent reviewers. Discrepancies were resolved through the intervention of another reviewer.

Quality assessment

The 10-item Drummond checklist was used to assess the methodological quality of the included studies (23). Details of the quality assessment are available in Supplementary Table 2. The Drummond checklist provides a global assessment of the quality of evidence, but it did not form the basis for accepting or rejecting articles.

Data Extraction

The researchers reviewed the selected full texts for eligibility and extracted the required data. For each publication, the following information was retrieved:

- Study characteristics: Publication year, Country and setting, Sample size, Prevalence of osteoporosis, Recruitment design, Prospective economic evaluation;
- Technique characteristics: Types of devices; Sites of application;
- Screening strategies with DXA and QUS, specifying QUS parameters and thresholds for women who required a DXA measurement for accurate diagnosis;
- Economic evaluation characteristics: Types of costs, Currencies and financial years; Cost breakdowns (i.e., the systematic process of identifying the individual elements that composed the unit costs of the QUS and DXA tests);
- DXA and QUS test results: Number of osteoporotic subjects accurately diagnosed (true positives); QUS sensitivity and specificity when available.

Economic analysis

Our review was conducted to analyse the cost per true positive case of two different strategies for osteoporosis screening and their incremental cost-effectiveness.

Cost per true positive case was calculated as total cost divided by number of true positive cases detected with the two different approaches: (1) the total cost per osteoporotic subject based on DXA measurement alone, i.e., without a QUS screen and (2) the total cost per osteoporotic subject identified by QUS+DXA, i.e., using QUS as a screen. This cost was the sum of the total cost of performing the QUS test in all subjects and the cost of performing additional DXA testing in those women who were positively detected with QUS.

For DXA, osteoporosis is defined by the WHO as a BMD that is 2.5 standard deviations or more below the mean peak bone mass in healthy young adults (a T-score $\leq$ -2,5) (24).

Incremental cost-effectiveness was calculated as the extra cost needed to generate each additional true positive result.

To compare the costs per true positive case, the current costs of the DXA and QUS tests were adjusted to Euro currency and inflation (base year 2006, i.e., the last year in the published studies used to estimate the test costs) (25,26) and exchange rates (27).

Cut-off values were calculated that indicated the level below which, based on the ratio unit cost of the QUS test and the unit cost of the DXA test, a true positive case diagnosed by the QUS+DXA technique was more cost-effective than was a true positive case detected by DXA alone.

Results

Overall, 136 publications were found. After the titles and abstracts were read, 85 publications were excluded as irrelevant (lack of pertinence or duplicates) or as reviews.

Of the remaining 51 publications, 40 were excluded because they did not meet the inclusion criteria. Finally, a total of 11 publications, 7 journal articles and 4 reports were included in our review (Fig. 1). The four reports provided the basis for the discussion.

The quality of the journal articles was good. Each article fulfilled six to ten items on the Drummond checklist (Supplementary Table 2).

The characteristics of the studies are illustrated in Table 1.

All journal articles were cohort studies, and their analyses were performed from the perspective of the third-party payer. The women's ages ranged from 40 to 70. The prevalence of osteoporosis ranged between 7.85% and 57.70%.

Three studies used mcCue CUBA Clinical (Mc Cue Plc, Winchester, UK), two used Wolkers Sonix UBA575 (Walker Sonics Inc. Worcester, MA) and one used the Sahara Clinical Bone Sonometer (Hologic Inc., Bedford, MA) for QUS test (Table 1). Three studies measured the broadband ultrasound attenuation (BUA) of the right calcaneus, two of the left calcaneus and two of both calcanea.

In terms of strategy, five studies adopted one QUS threshold value, and two studies used different QUS threshold values to identify women who needed a DXA measurement for an accurate diagnosis.

Five studies used BUA measurements as the QUS parameter and two used T-scores (Table 1).

Five studies reported the real cost of the DXA and QUS tests; two studies reported the charges or estimated costs (Table 1). Two of the 7 studies gave a breakdown of different cost items (Table 1).

All studies reported that the DXA test was costlier than the QUS test, with some stating that the cost of the DXA test was nine-fold or eight-fold higher (28-31) and others reporting that it was two-fold (32), three-fold (31,32), four-fold (33) or five-fold higher (34) (Table 2).

According to Langton (28,29) and Marin (30), the cost per true positive case diagnosed by DXA was higher than that for diagnosis by QUS+DXA. Kraemer (33), however, estimated that the cost per osteoporotic subject identified by DXA alone was less than the cost per osteoporotic subject identified by QUS+DXA. In contrast, three studies (31,32,34) reported that the cost per osteoporotic subject identified by DXA alone was higher or lower than that of QUS+DXA (Figure 2).

For all of the studies, a cut-off value was calculated that indicated under what ratio unit cost for the QUS and DXA tests a case diagnosed by combining QUS+DXA was more cost-effective than was a case diagnosed by DXA alone. Depending on the study, cases diagnosed by QUS+DXA were cost-effective as long as the cost of the QUS test was between 7% and 41% of the cost of the DXA test (for each study, the cut-off values were as follows: 41% in the studies by Langton, 30% in Sim 2000 and 36% in Sim 2005, 14% in Marin, 7–22% in Kraemer, and 13–32% in Hiligsmann) (Table 2).

The incremental cost to diagnose one more case ranged approximately between 30 and 2,000 Euros (Table 2).

In three cases, there were incremental savings associated with diagnosing each additional case: QUS 80 Db/MHz and QUS 85 Db/MHz, in Kraemer (33) with a QUS T-score = 0 in Hiligsmann (34) (Table 2).

Discussion

The policy question we posed for this study was whether DXA alone was more cost-effective for identifying postmenopausal women with osteoporosis than was a two-step procedure using QUS plus DXA. In a previous review in 2008, Schousboe (35) concluded that on balance, the cost-effectiveness studies of heel ultrasounds did not make a convincing case that heel ultrasounds should be employed in places where central DXA was available. In his review, Schousboe did not compare cost per true positive case detected by DXA alone with cost for QUS+DXA; he did not analyse variables such as, e.g., sensitivity and specificity or type of costs. In contrast, in our review, we made a comparison and took into account a number of variables.

In any event, the results of our review did not allow for definitive conclusions about the better technique for diagnosing osteoporosis, most likely because of the lack of homogeneity among the studies. One of the difficulties we encountered in comparing the studies was the different QUS devices used. They differ substantially with respect to the algorithms they used, the parameters they measured, and the strength of the empirical evidence supporting their use, among other aspects (36,37). Another difficulty we encountered in comparing the studies was the different ways the QUS parameter was measured: BUA with different Db/MHz values or a QUS T-score (0.0, -.05, -1.0, -1.5, -2.0, -2.5). According to the INAHTA report (37), these parameter measurements are not directly comparable.

In addition, there are different sites for the device applications. Only Sim (31, 32) measured the BUA of the right and left calcanea and showed that QUS sensitivity increased with the use of the left calcaneus in comparison with that of the right calcaneus.

Furthermore, there were differences concerning the costs for true positive cases.

Our review shows the inhomogeneous results likely caused by (i) the types of costs used to determine the values of the QUS and DXA tests, (ii) the items included in the evaluations of the DXA and QUS test costs, (iii) the different sensitivities and specificities of the QUS tests and (iv) the ages of the screened populations and their prevalence of osteoporosis.

Indeed and about point:

(i) Kraemer (33) found the cost per QUS test and DXA test to be higher than did any other study, perhaps because he used Medicare reimbursement rates rather than real costs; (ii) the costs of the DXA test in Sim 2005 (32) consisted of staff salaries (technologist, doctors, nurse, clerk, training and development), equipment (depreciation over 7 years, maintenance, consumables interest) and assumed overhead of 20% or more. In Marin (30), costs consisted of salaries for the technician and the doctors, equipment and maintenance and assumed overhead of 10%. The different resources identified, enumerated and valued determined that the DXA test in Sim 2005 (32) was more expensive than that used in Marin (30); (iii) in Hiligsmann (34), the cost per true positive case detected by QUS+DXA was higher than the cost for diagnosis by DXA alone when the QUS T-score = -2.5 or = 0.0. In the other situations, true positive cases diagnosed by QUS+DXA were less expensive than were those identified by DXA alone. This could be explained by the range in the QUS sensitivity and specificity, between 33% and 93% and between 24% and 93%, respectively, so that the cost per true positive case diagnosed by QUS+DXA was higher when there were many false positives or there was low QUS sensitivity; (iv) a significant factor in the cost per true positive case is the prevalence of osteoporosis at the different ages. Langton 1999 (29) showed that for women aged 50–55, the prevalence of osteoporosis was only 7.85%, whereas Langton 1997 (28) found that the prevalence for women aged 60–69 was 24.3%. As the prevalence of osteoporosis within a population increases, the total screening cost is divided over a large number of osteoporotic subjects and the cost per subject identified decreases. Hence, the cost per true positive case diagnosed by DXA in Langton 1999 (29) was 967.83 Euros for the 50–54 year cohort, falling to 312.52 Euros for the 60–69 year cohort in Langton 1997 (28).

Most of the authors suggested an extra cost needed to generate each additional true positive result using DXA compared with QUS+DXA because of the higher DXA test unit cost and the low QUS sensitivity. These results confirmed Nayak's (18) conclusions. Nayak demonstrated that different osteoporosis screening methods were effective and that there were incremental costs for DXA

screening per additional identified case. In contrast, Kraemer (33) suggested cost savings per additional true case of osteoporosis diagnosed by DXA when the QUS parameter was 80 and 85 Db/MHz and the sensitivity was 91% and 95% for QUS respectively, similar to Hiligsmann (34), with a QUS T-score=0 and a sensitivity of 93%.

In conclusion, our review aimed to assess the available evidence on the cost-effectiveness of different techniques for diagnosing osteoporosis in postmenopausal women.

Although there is some evidence that screening is effective in identifying postmenopausal women with osteoporosis, our results suggest that the role of QUS in the diagnosis of osteoporosis remains unclear (39,40) and show that from the perspective of the third-party payer, QUS may be useful as a pre-screening tool for osteoporosis if the cost ratio between QUS test and DXA test unit costs is below a specified cut-off value (expressed in %) and if the QUS sensitivity and specificity are high.

To arrive at a definitive conclusion of whether DXA alone is more cost-effective for identifying postmenopausal women with osteoporosis than is a two-step procedure with QUS plus DXA and is in line with the INAHTA report (41), our results underscored that homogenous cost-effectiveness studies are needed to elucidate the question as to which technique is less costly and more effective in the identification of patients with osteoporosis. In this way, some of the studies' biases could be overtaken, e.g., conclusions cannot be extended to women younger or older than the target group being examined; costs and resource use that were not adequately reported. The problem is relevant because the experts agreed that in the future, fracture prediction will change with the use of the more complex FRAX® system, which integrates both DXA and QUS+DXA data.

In addition, the problem is significant because the evidence for which is the most cost-effective—DXA only or QUS + DXA—is important for policy makers, who have then to combine these results with other information about possible interventions for treating osteoporosis and actual health outcomes, such as osteoporosis-related fracture reduction.

In fact, because there is evidence that fractures and their complications are the relevant clinical sequelae of osteoporosis; that osteoporosis-related fractures create a heavy economic burden; and that patients with osteoporosis reduced their fracture risk with pharmacotherapy, a comprehensive approach to diagnosing and managing osteoporosis is recommended to decision makers. This approach should take into account the cost of the screening programme but also the cost of the resources necessary to treat true positive cases and the benefits of these resources to health outcomes.

In addition to being homogeneous, studies on cost-effectiveness must be conducted with greater methodological rigor because health care decision makers need to be sure that the evidence on efficiency is reliable and can be applied to their own situations.

In this review, the process of critically appraising health economic evaluation studies assisted by Drummond checklists showed that the quality of published health economic evaluations varies. Some studies did not provide sufficient evidence for decision makers in the countries in which they were conducted: e.g., Langton (28,29) in the UK and Kramer (33) in the United States did not provide details of the costs of the two tests and did not justify the types of costs used, and Hiligsmann (34) in Belgium did not employ a sensitivity analysis. Thus, decision makers cannot judge if the findings are applicable to their health service or social insurance.

References

1. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. Osteoporosis prevention, diagnosis, and therapy. JAMA 2001; 285(6):785-95.
2. Gambacciani M, De Aloysio D, Elia D, et al. Quantitative ultrasound (QUS) of bone in the management of postmenopausal women. Maturitas. 2004;20;47(2):139-49.
3. National Osteoporosis Foundation. About Osteoporosis: Fast Facts. 2005. Washington, DC: National Osteoporosis Foundation.
4. Nelson HD, Helfand M, Woolf SH, et al. Screening for postmenopausal osteoporosis: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002;137:529-41.

5. Cooper C, Campion G, Melton LJ 3rd. Hip fractures in the elderly: a world-wide projection. Osteoporos Int. 1992;2(6):285-9.
6. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrance of osteoporotic fractures. BMJ. 1996;312:1254-1259.
7. McCloskey EV, Johansson H, Oden A, et al. FRom relative risk to absolute fracture risk calculation: the FRAX® in osteoporosis. Curr Osteoporos Rep. 2009;7:77-83.
8. Kanis JA, McCloskey EV, Johansson H, et al. development and use of FRAX® in osteoporosis. Osteoporosis Int. 2010;21(Suppl. 2): S407-S413.
9. Lakatos Lakatos P, Balogh A, Czerwinski E, et al. New considerations on the management of osteoporosis in Central and Eastern Europe (CEE): summary of the "3rd Summit on Osteoporosis-CEE", November 2009, Budapest, Hungary. Arch Osteoporos. 2011;6(1-2):1-12.
10. WHO Study Group. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organ Tech Rep Ser. 1994;843:1–129.
11. Genant HK, Faulkner KG, Gluer CC, et al. Bone densitometry: current assessment. Osteoporosis Int. 1993;51:91-97.
12. Lewiecki EM, Richmond B, Miller PD. Uses and misuses of quantitative ultrasonography in managing osteoporosis. Cleve Clin J Med. 2006;73(8):742-6,749-52.
13. Nayak S, Olkin I, Liu H, et al. Meta-analysis: accuracy of quantitative ultrasound for identifying patients with osteoporosis. Ann Intern Med. 2006;144(11):832-41.
14. Faulkner KG, von Stetten E, Miller P. Discordance in patient classification using T-scores. J Clin Densitom 1999;2:343–350
15. MacLaughlin EJ, MacLaughlin AA, Snella KA, et al. Osteoporosis Screening and Education Community Pharmacies using a team approach. Pharmacotherapy 2005;25(3):379-386.
16. Marín F, González-Macías J, Díez-Pérez A, et al. Relationship between bone quantitative ultrasound and fractures: a meta-analysis. J Bone Miner Res. 2006;21(7):1126-35.
17. Schousboe JT, Gourlay ML. Comparative effectiveness and cost-effectiveness of strategies to screen for osteoporosis in postmenopausal women. Ann Intern Med. 2011; 155(11):788-9
18. Nayak S, Roberts M, Greenspan S. Cost-effectiveness of different screening strategies of osteoporosis in postmenopausal women. Ann Intern Med 2011;155(11):751-61.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;21;6(7).

20. Gianino MM, Galzerano M, Tizzani A, et al. Critical issues in current comparative and cost analyses between retropubic and robotic radical prostatectomy. BJU Int. 2008;101(1):2-3.
21. Gianino MM, Galzerano M, Minniti D, et al. A comparative costs analysis of brachytherapy and radical retropubic prostatectomy therapies for clinically localized prostate cancer.Int J Technol Assess Health Care. 2009;25(3):411-4.
22. Leib ES, Lewiecki EM, Binkley N, et al. Official positions of the International Society for Clinical Densitometry. J Clin Densitom 2004;7:1-6.
23. Drummond M et al. Methods for the economic evaluation of health care programmes. 2nd ed. Oxford. Oxford University Press. 1997
24. WHO Study Group. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organ Tech Rep Ser. 1994;843:1–129.
25. TradingEconomic.com- Economic Data for 196 Countries. 2012. **http://www.tradingeconomics.com**
26. UK National Statistics Publication Hub. 2012. Available at: **http://www.statistics.gov.uk**
27. BCE: The European Central Bank. 2012. Available at: **www.ecb.int**
28. Langton CM, Ballard PA, Langton DK, et al. Maximising the cost-effectiveness of BMD referral for DXA using ultrasound as a selective population pre-screen. Technol Health Care 1997;5(3):235-41.
29. Langton CM, Langton DK, Beardsworth SA. Comparison of accuracy and cost-effectiveness of clinical criteria and BUA for referral for BMD assessment by DXA in osteoporotic and osteopenic perimenopausal subjects. Technol Health Care 1999;7(5):319-30.
30. Marín F, López-Bastida J, Díez-Pérez A, et al. Bone mineral density referral for dual-energy X-ray absorptiometry using quantitative ultrasound as a prescreening tool in postmenopausal women from the general population: a cost-effectiveness analysis. Calcif Tissue Int. 2004;74(3):277-83.
31. Sim MF, Stone M, Johansen A, et al. Cost-effectiveness analysis of BMD referral for DXA using ultrasound as a selective pre-screen in a group of women with low trauma Colles' fractures. Technol Health Care 2000;8(5):277-84.
32. Sim MF, Stone MD, Phillips CJ, et al. Cost-effectiveness analysis of using quantitative ultrasound as a selective pre-screen for bone densitometry. Technol Health Care 2005;13(2):75-85.

33. Kraemer DF, Nelson HD, Bauer DC, et al. Economic comparison of diagnostic approaches for evaluating osteoporosis in older women. Osteoporos Int 2006;17(1):68-76.
34. Hiligsmann M, Ethgen O, Bruyere O, et al. An economic evaluation of quantitative ultrasonometry as pre-screening test for the identification of patients with osteoporosis. Disease Management and Health Outcomes 2008;16(6):429-438.
35. Schousboe JT. Cost effectiveness of screen-and-treat strategies for low bone mineral density: how do we screen, who do we screen and who do we treat? Appl Health Econ Health Policy. 2008;6(1):1-18.
36. Langton C, Njeh c. The measurement of broadband ultrasonic attenuation in cancellous bone. A review of the science and technology. Ieee Transaction on ultrasonics, ferroelectrics and frequency control. 2008;55(7):1546-1554.
37. Njeh CF, Hans D, Fuerst T, et al. Quantitative ultrasound: Assessment of osteoporosis and bone status. 1999 London: Martin Dunitz, Ltd.
38. Alton V. Osteoporosis Prevention, Diagnosis, and Treatment. INAHTA Brief issue: 2003/70 (available from http://www.sbu.se/upload/Publikationer/Content1/1/Eng_Osteoporos.pdf consulted 14 Dec 2012)
39. Homik J, Hailey D. Quantitative ultrasound for bone density measurement. Alberta Heritage Foundation for Medical Research, 1998 (Project page URL http://www.ihe.ca/hta/hta_ unit.html, bibliographic data drawn from: http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=31999008037 consulted 14 Dec 2012)
40. Institute for Clinical Systems Improvement Densitometry as a diagnostic tool for the identification and treatment of osteoporosis in women. Bloomington MN: Institute for Clinical Systems Improvement; 2000 (Project page URL http://www.icsi.org/index.asp, bibliographic data drawn from: http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=32003000546 consulted 14 dec 2012)
41. Dunfield L, Mierzwinski-Urban M, Hodgson A, et al. Diagnostic Performance and Cost-Effectiveness of Technologies to Measure Bone Mineral Density in Postmenopausal Women. [Technologyreport number 94]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2007

Figure legends

Figure 1. Study Flow Diagram – steps for selecting the studies for inclusion in this review.

Figure 2. Cost per true positive case with DXA* and QUS+DXA in 2006 Euros.**

Table 1. The characteristics of the studies

Study and publication year	Country (Setting)	Sample size	Recruitment design	Prevalence of osteoporosis	Perspective of economic evaluation	Type of device	Site of application	Strategies	Types of costs, currency and financial year	Cost breakdown
Langton (1997) [28]	UK (Centre for Metabolic Bone Disease, Hull)	107 women aged 60–69, mean age, 64.2±2.8	NA	24.30%	National Health Services in UK	DXA (pencil-beam Lunar DPX-L); QUS (CUBA clinical II scanner, McCue Ultrasonic, Winchester, UK)	Right femoral neck or lumbar spine for DXA; Left calcaneus for QUS	1 DXA for all 2. QUS for all and DXA for those with BUA score QUS =60 Db/MHz	Real cost*, £, 1997	No
Langton (1999) [29]	UK (Centre for Metabolic Bone disease,	599 women aged 50–	Cohort of women was provided from	7.85%	National Health Services in	DXA (pencil-beam Lunar DPX-L);	Right femoral neck or	1 DXA for all 2. QUS	Real cost*, £, 1997	No

	Hull)	54, mean age, 52.18±1.35	the data set of a different study		UK	QUS (Walker Sonix UBA575)	lumbar spine for DXA; Right calcaneus for QUS	for all and DXA for those with BUA score QUS =75 Db/MHz		
Sim (2000) [31]	UK (Cardiff Royal Infirmary)	46 women aged 50–80, mean age 69	Cohort of consecutive women who presented at the Accident and Emergency Department at the Cardiff Royal Infirmary	58.70%	National Health Services in UK	DXA (Hologic QRD 1000W, Hologic Inc., Waltham, MA, USA); QUS (CUBA clinical II scanner, McCue Ultrasonic, Winchester, UK)	Total hip or lumbar spine for DXA; Left calcaneus for QUS	1 DXA for all 2. QUS for all and DXA for those with BUA score QUS =60 Db/MHz	Real cost* Estimated cost**, £, 1997	No
Marin (2004) [30]	Spain (Three primary care centres in metropolitan	267 women aged ≥ 65, mean age	Cohort of women without neoplastic or	55.81%	National Health Services in Spain,	DXA (Hologic QRD 4500SLTM, Hologic Inc.,	Femoral neck for DXA; Heel for	1 DXA for all 2. QUS for all and	Real cost*, €, 2001	Yes

	Barcelona)	72.3±5.3	metabolic bone disease who were attended to for any medical reason in three primary care centres (non-probabilistic sampling of consecutive cases)			Waltham, MA, USA), QUS (Sahara Clinical Bone Sonometer, Hologic Inc., Bedford).	QUS	DXA for those with BMD QUS T-score <0.5 to >-2.5			
Sim (2005) [32]	UK (Cardiff open access bone densitometry service)	115 women aged 40–80, mean age 69	All consenting women referred by their GPs via the Cardiff open access bone densitometry service	46.09%	National Health Services in UK	DXA (Hologic QRD 1000W, Hologic Inc., Waltham, MA, USA); QUS (CUBA clinical II scanner, McCue Ultrasonic,	Total hip or lumbar spine for DXA; Right calcaneus for QUS	1 DXA for all 2. QUS for all and DXA for those with BUA score QUS =60	Real cost*, £, 2002	Yes	

					Winchester, UK)		Db/MHz		
Kramer (2006) [33]	USA (NA)	5993 women aged ≥ 70	Only women with both DXA and QUS measurements within Study of Osteoporotic Fractures (SOF) cohort	28.10%	Medicare	DXA (QDR 1000 instrument, Hologic, Bedford, Mass.), QUS (UBA 575 instrument Walker-Sonix, Worcester, Mass.).	Femoral neck for DXA; Calcaneus for QUS	1 DXA for all 2. QUS for all and DXA for those with BUA score QUS =50; 55; 60; 65; 70; 75; 80 e 85 Db/MHz	Charge, $, 2000	No
Hiligsmann (2008) [34]	Belgium (NA)	1000 women aged 50-80	No recruitment but hypothetical cohort	19.93%	Social Insurance	NA	NA for DXA; Calcaneus for QUS	1 DXA for all 2. QUS for all and DXA for those with	Cost estimated**, €, 2006	No

	BMD QUS T-score =0;=-0.5;=-1; =-1.5; =-2;≤ -2.5

*Cost estimated= when resources that required alternative methods (QUS and DXA) were proposed by institutions or the literature.

**Real cost= when actual resources that required alternative methods (QUS and DXA) were identified, enumerated and valued.

Table 2. Sensitivity, Specificity, Unit cost of the QUS test and DXA test, incremental cost and cut-off value (2006 euros)

Study and publication year	Parameter and Thresholds in QUS for osteoporosis	Unit cost DXA test (euros)	Unit cost QUS test (euros)	QUS sensitivity	QUS specificity	Incremental cost-effectiveness of DXA alone vs QUS+DXA (euros) [1]	Cut-off value[2]
Langton (1997) [28]	BUA score QUS =60 Db/MHz	75.94 [a]	8.18 [a]	73%	81%	666.84	<41%
Langton (1999) [29]	BUA score QUS =75 Db/MHz	75.94 [a]	8.18 [a]	73%	73%	2052.96	<41%
Sim (2000) [31]	BUA score QUS = 60 Db/MHz	75.94 [a]	25.31 [ab]	93%	84%	101.25	<30%
		75.94 [a]	8.18 [b]	93%	84%	495.21	<29%
Marin (2004) [30]	BMD QUS T-score ≤ -2.5	14.77 [b]	1.84 [b]	97%	94%	126.51	<14%
Sim (2005) [32]	BUA score QUS = 60 Db/MHz	50.47 [c *]	25.99 [c]	81%	89%	29.20	<36%
		69.59 [c**]	25.99 [c]	81%	89%	153.46	<36%
Kramer (2006) [33]	BUA score QUS 50 Db/MHz	128.78 [d]	32.92 [d]	34% [e]	90% [e]	400.35	<22%
	BUA score QUS	128.78 [d]	32.92 [d]	47% [e]	82% [e]	417.97	<17%

	55 Db/MHz						
	BUA score QUS 60 Db/MHz	128.78 [de]	32.92 [d]	60% [e]	71% [e]	423.21	<20%
	BUA score QUS S 65 Db/MHz	128.78 [d]	32.92 [d]	71% [e]	59% [e]	396.32	<21%
	BUA score QUS 70 Db/MHz	128.78 [d]	32.92 [d]	79% [e]	47% [e]	304.18	<10%
	BUA score QUS 75 Db/MHz	128.78 [d]	32.92 [d]	86% [f]	35% [e]	106.28	<18%
	BUA score QUS 80 Db/MHz	128.78 [d]	32.92 [d]	91% [e]	24% [e]	-314.64	<7%
	BUA score QUS 85 Db/MHz	128.78 [d]	32.92 [d]	95% [e]	16% [e]	-1070.04	<15%
Hiligsmann (2008) [34]	BMD QUS T-score ≤ -2.5	47.00	10.00	33%	93%	234.52	<19%
	BMD QUS T-score = -2	47.00	10.00	49%	86%	266.90	<28
	BMD QUS T-score =-1.5	47.00	10.00	66%	74%	312.25	<32%
	BMD QUS T-score =-1	47.00	10.00	79%	58%	330.88	<28%
	BMD QUS T-	47.00	10.00	88%	39%	248.39	<21%

score = -0,5						
BMD QUS T-score =-0	47.00	10.00	93%	24%	-23.70	<13%

[1] Incremental cost –effectiveness is calculated as the extra cost needed to generated each additional true positive result

[2] A cut-off value was calculated and indicated under what ratio unit cost for the QUS and DXA tests a case diagnosed by combining QUS+DXA was more cost-effective than was a case diagnosed by DXA alone

[a] The values in € 2006 are obtained adjusting for an inflation rate of 1.1332 and for an exchange rate £/€ of 1.,4892

[b] The values in € 2006 are obtained adjusting for an inflation rate of 1.1095

[c] The values in € 2006 are obtained adjusting for an inflation rate of 1.1035 and for an exchange rate £/€ of 1.4892

[d] The values in € 2006 are obtained adjusting for an inflation rate of 1.1710 and for an exchange rate $/€ of 0.,8268

[e] Our own calculations

[*] DXA fan beam

[**] DXA pencil beam

Supplementary Table 1. Search strategies in 2012

Database	Search strings
Pubmed	("Bone Density"[Mesh] OR "Bone and Bones"[Mesh]) AND ("Quantitative ultrasonometry" OR "Quantitative ultrasound" OR QUS) AND ("Dual-energy X-ray absorptiometry" OR DXA OR DEXA) AND (econom* OR cost* OR "Economics"[Mesh])
Pubmed	("Quantitative ultrasonometry" OR "Quantitative ultrasound" OR QUS) AND ("Dual-energy X-ray absorptiometry" OR DXA OR DEXA) AND (econom* OR cost* OR "Economics"[Mesh])
International Network of Agencies for Health Technology Assessment INAHTA Health Evidence Network (HEN) National Institute for Health Research (NIHR) Health Technology Assessment programme National Health Service (NHS) Economic Evaluation Database Research Papers in Economics (RePEc)	QUS AND DXA OR DEXA
WEB OF SCIENCE SCOPUS*	Topic=(qus) AND Topic=(dxa) AND Topic=(econom*)

*SCOPUS was searched with a different strategy from that used for Pubmed, which allowed for more detailed and sophisticated searches.

Supplementary Table 2. Quality assessment with the 10-item Drummond checklist

Check list										
1.Was a well-defined question posed in an answerable form?										
2.Was a comprehensive description of the competing alternatives given (i.e., could you tell who did what to whom, where, and how often)?										
3.Was the effectiveness of the programme or services established?										
4.Were all of the important and relevant costs and consequences for each alternative identified?										
5.Were the costs and consequences measured accurately in appropriate physical units (e.g., hours of nursing time, number of physician visits, lost work days, gained life years)?										
6.Were the cost and consequences valued credibly?										
7.Were the costs and consequences adjusted for differential timing?										
8.Was an incremental analysis of the costs and consequences of alternatives performed?										
9.Was allowance made for uncertainty in the estimates of the costs and consequences?										
10.Did the presentation and discussion of the study results include all issues of concern to the users?										
Study and publication year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Langton (1997) [28]	√	√	√	√	**0**	**0**	√	**0**	**0**	√
Langton (1999) [29]	√	√	√	√	**0**	**0**	√	**0**	**0**	√
Sim (2000) [31]	√	√	√	√	**0**	**0**	√	**0**	√	√
Marin (2004) [30]	√	√	√	√	√	√	√	√	√	√
Sim (2005) [32]	√	√	√	√	**0**	√	√	**0**	**0**	√
Kramer (2006) [33]	√	√	√	√	**0**	**0**	√	**0**	√	√
Hiligsmann (2008) [34]	√	√	√	√	**0**	**0**	√	√	**0**	√

Yes =√ ; No=0

Figure 1. Study Flow Diagram – steps of a selection of studies for inclusion in this review.

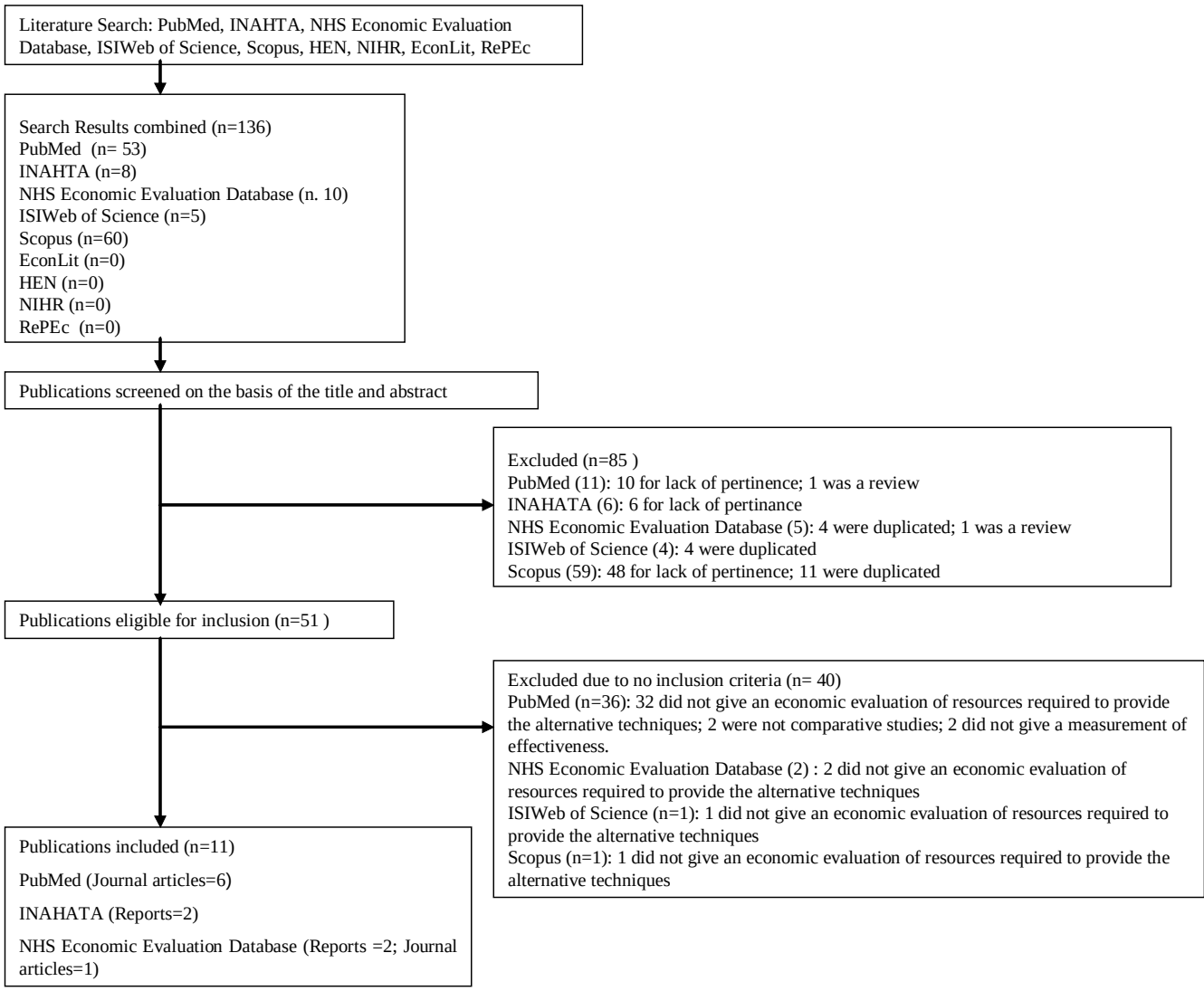
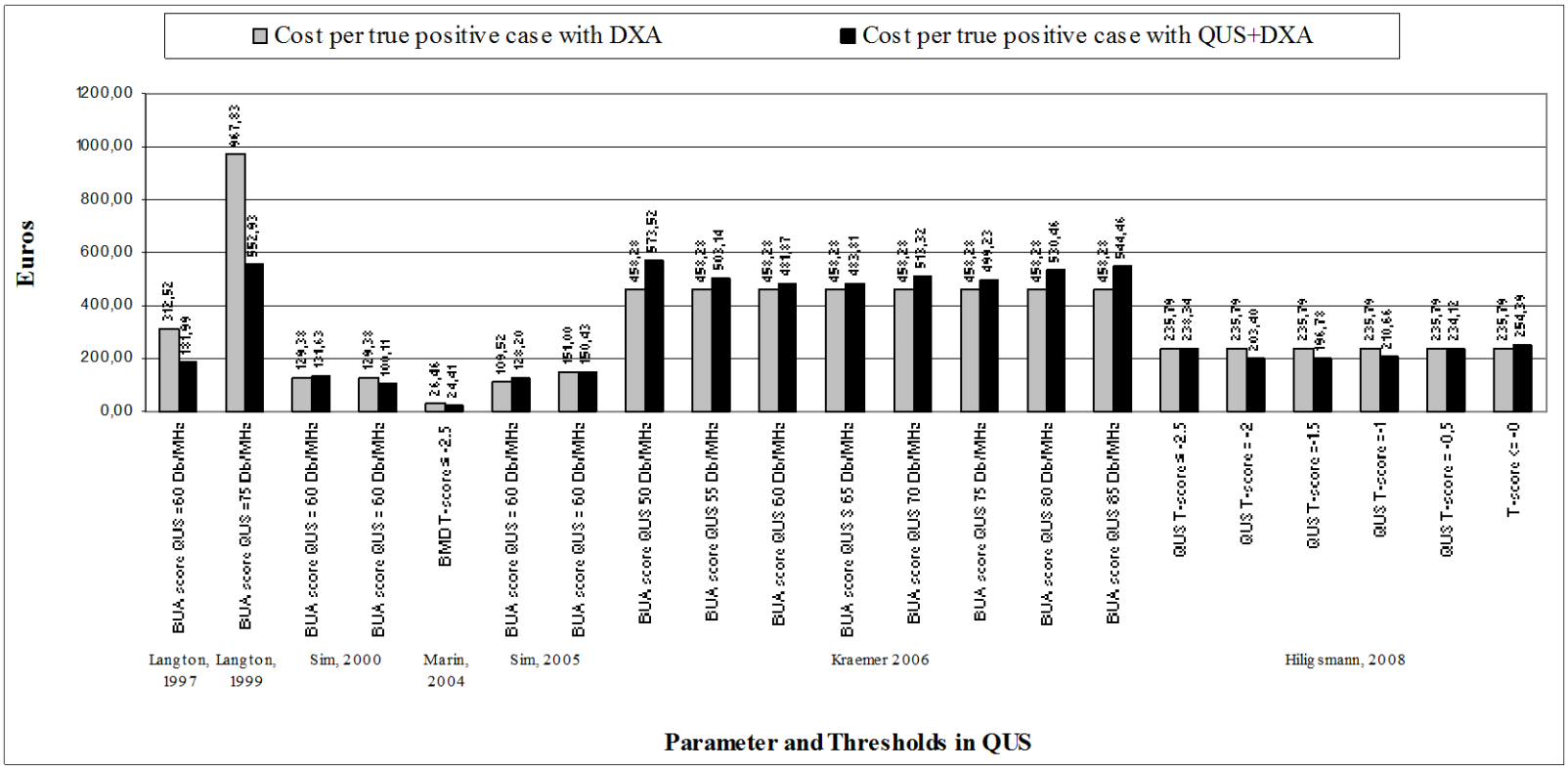


Figure 2. Cost per true positive case with DXA* and QUS+DXA in 2006 euros.**



*Cost per true positive case was calculated as the total cost divided by the number of true positive cases detected with DXA

**Cost per true positive case was calculated as the total cost divided by the number of true positive cases detected with QUS+DXA. The total cost is the sum of the total cost of performing the QUS test on all subjects and the cost of performing additional DXA on those women who were positively detected with QUS.