Arch Osteoporos
DOI 10.1007/s11657-010-0045-5

REVIEW

An evaluation of the NICE guidance for the prevention of osteoporotic fragility fractures in postmenopausal women

John A. Kanis · Eugene V. McCloskey · Bengt Jonsson · Alun Cooper · Oskar Ström · Fredrik Borgström

Received: 21 June 2010 / Accepted: 24 September 2010

Abstract
Summary The National Institute for Health and Clinical Excellence (NICE) in the UK issued guidance based on a health economic assessment of interventions for the primary and secondary prevention of osteoporosis. The recommendations in the guidance are unworkable in clinical practice and the foundation on which they are based is insecure.
Introduction The NICE in the UK recently issued final appraisal documents on the health economic assessment of interventions for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. The majority of interventions were considered to be cost-ineffective except at very low *T* scores for bone mineral density (BMD). Concerns have been raised with respect to the construct and assumptions that populate the model used by NICE and the feasibility of implementing the subsequent guidance.
Results The application of the NICE guidance to primary care is problematic. Intervention thresholds are based on a complex array that includes the agent to be used, age, *T* scores and the presence of different categories of risk factors. Alendronate is the first-line treatment, but women who cannot take or tolerate alendronate may have to wait till their *T* score deteriorates before they qualify for treatment. The guidance takes no account of women with a *T* score>−2.5 SD, of glucocorticoid-induced disease or of men. Newer interventions, such as ibandronate and zoledronic acid, are not included. The development of guidelines by the National Osteoporosis Guideline Group (NOGG) avoids many of these problems and unlike the NICE guidance, can be used with FRAX®, the WHO-supported fracture risk assessment tool. NOGG provides intervention thresholds based on fracture probabilities computed from clinical risk factors for fracture with or without information on BMD that are readily accessed by primary care physicians for the assessment of all postmenopausal women and men over the age of 50 years. The NICE guidance is based on a health economic assessment of several interventions. The model used to assess cost-effectiveness is based on Gaussian regression functions which were derived from an individual state transition model. Since the source individual state transition model is not available, the Gaussian functions cannot be evaluated. Moreover, neither the internal nor external validity of the model is established, and the model is not accessible for such an evaluation. Although the NICE model incorporates the clinical risk factors (CRFs) used in FRAX, it neglects the impact of CRFs on the death hazards giving estimates of fracture probability that differ from those using FRAX®.

J. A. Kanis (✉)
WHO Collaborating Centre for Metabolic Bone Diseases,
University of Sheffield Medical School,
Beech Hill Road,
Sheffield S10 2RX, UK
e-mail: w.j.Pontefract@shef.ac.uk

E. V. McCloskey
Academic Unit of Bone Metabolism, Northern General Hospital,
Sheffield, UK

O. Ström · F. Borgström
i3 Innovus,
Stockholm, Sweden

B. Jonsson
Stockholm School of Economics,
Stockholm, Sweden

A. Cooper
Crawley Fracture Liaison Service,
Crawley, UK

O. Ström · F. Borgström
Medical Management Centre, Karolinska Institute,
Stockholm, Sweden

The estimates of cost-effectiveness differ from reference models for reasons that relate in part to the model construct and in particular to the assumptions used to populate the model.

Conclusions The guidance provided by NICE is cumbersome and cannot be readily used in the setting of primary care. The model on which the guidance is based is opaque. The authors do not support the view of NICE that there are no issues which cause it to doubt the validity of the model or that raise justifiable doubts about the appropriateness of the use of the model to inform its guidance.

Keywords FRAX · Economic models · Fracture probability · Clinical risk factors · Intervention thresholds

Abbreviations

BMI	Body mass index (computed as kg/m^2)
BMD	Bone mineral density (in this report at the femoral neck measured by dual energy X-ray absorptiometry)
CRFs	Clinical risk factors
DSU	Decision support unit
FAD	Final appraisal document
FRAX®	Algorithms that assess the probability of fracture related to any combination of clinical risk factors with or without BMD
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
NICE	National Institute for Health and Clinical Excellence
NOGG	National Osteoporosis Guideline Group
QALY	Quality-adjusted life years
RCP	Royal College of Physicians, London
ScHARR	School of Health and Related Research, University of Sheffield, UK
T score	The deviation in SD units of measured BMD from the mean of the young adult female reference range
WHO	World Health Organization
WTP	Willingness to pay

Introduction

The clinical consequences of osteoporosis reside in the fractures that arise, particularly hip fracture which accounts for the major direct costs. In 1990, the number of osteoporotic fractures estimated in Europe was 2.7 million, with a direct cost in 2004 of €36 billion (£24.5 billion), of which €24.3 (£16.6) billion were accounted for by hip fracture. Costs are expected to rise to €76.8 (£52.4) billion by the year 2050 [1] because of the increasing numbers of the elderly. The financial cost to the UK is estimated at over £2 billion yearly for postmenopausal women [2].

Against this background of the burden of osteoporosis, there has been an increase in the number of agents available that have been shown in well-designed studies to decrease the risk of fractures [3, 4]. Recommendations concerning the use of these agents in the UK and several other countries have been placed in a health economic setting in order to justify resource allocation and form the basis for the development of clinical guidelines. The agency responsible for this in the UK is the National Institute for Health and Clinical Excellence (NICE).

The history of NICE guidance for osteoporosis is long and complex. The appraisal began in 2002 under the management of an Appraisal Committee with the provision of the scope and guidance for secondary prevention of fracture produced in January 2005 [5]. Subsequently, the scope was broadened to include strontium ranelate and new guidance for both primary and secondary prevention was issued in June 2007 [6, 7]. Following a successful appeal against the guidance, new guidance was released in October 2008 [8, 9]. This was brought to judicial review in January 2009 [10, 11] and further appraisals were published in October 2009 but amended in January 2010 [12, 13]. A more recent appeal to the High Court in December, 2009 ruled in March 2010 that NICE had provided insufficient justification for its assumptions on the efficacy of strontium ranelate, and ordered that NICE issue new guidance in relation to strontium ranelate. The much protracted process has been due in part to concerns raised by the Guideline Development Group (GDG) and consultees about the model construct and the assumptions used to populate the model for these appraisals [14–17]. Several independent analyses have revealed major differences in cost-effectiveness measures from those published by NICE for alendronate, raloxifene, strontium ranelate and risedronate [16, 18–20]. Concern was intensified by the observation that recommendations for the use of alendronate had barely changed between 2005 and 2010 despite a sixfold reduction in price with the availability of generic alendronate [17]. This stability was achieved by alteration of some of the model assumptions in the absence of new evidence, so that the cost-effectiveness of alendronate remained unchanged despite its fall in price. Furthermore, these changes to the model had a negative impact on the cost-effectiveness of the other treatments under consideration.

A major concern was that NICE had not been transparent in providing access to the model used for the appraisal process. NICE argued that transparency was not possible because of the confidential nature of information provided to NICE for use in the model. In March of 2009, the High Court ruled that NICE had not acted reasonably in securing the release of the information under an appropriate confidentiality arrangement [21]. As a consequence, NICE

released a version of the model to interested consultees for comment and responses were submitted to NICE for evaluation and the final appraisal documents (FADs) of 2010 were subsequently issued. Transparency was not well served by the gagging by NICE of consultees who were not permitted to comment publicly on the output of the model. In addition, previous feedback from stakeholders, consultees and the public, which particularly addressed the changes in assumptions in the model over time, is no longer accessible from the NICE website and hence the record of the appraisal process is incomplete.

In parallel with each appraisal process, NICE appoints a GDG of experts in the relevant field, the task of which is to synthesise clinical *guidelines* that are based on the *guidance* developed by the appraisal process. In the case of osteoporosis, the GDG was appointed in 2002 and provided continuous feedback on all elements of the appraisals, prepared several systematic reviews of the evidence to inform guidance and developed its strategy for case finding. The relationship between the GDG and the Appraisal Committee was strained by the reluctance of the Appraisal Committee to heed the advice of the GDG in all but the most trivial issues. The GDG was suspended in 2007, its records remain confidential, its systematic reviews (with one exception [22]) are unpublished and there is no record of its guideline strategy. The void in guidelines occasioned by the lengthy process and the prospect of guidance that could not be supported by experts in the field gave rise to the National Osteoporosis Guideline Group that published its own guideline in 2008.

Against this background of obfuscation, the aim of this paper is to assemble the collective arguments that have been lost in the long evolution of the NICE appraisals and the impact that the process has had on the status of guideline development. To this end, the present paper reviews the model supplied by NICE, the assumptions used to populate the model that gave rise to the greatest concerns and the clinical difficulties that arose from the appraisals. These difficulties are first placed into the context of pre-existing guidelines in the UK [23].

The evolution of assessment guidelines in the UK

Royal College of Physicians

A number of case-finding strategies have been advocated to identify individuals at high risk for osteoporotic fractures [24]. Until recently, the most widely used guidelines in the UK were those provided by the Royal College of Physicians (RCP) [25, 26] that were based on those developed by the European Foundation for Osteoporosis (now the International Osteoporosis Foundation) [27]. Under this strategy, patients with clinical risk factors for fracture (CRFs) were identified and thereafter referred for testing with a bone mineral density (BMD) measurement. Treatment was recommended in patients with osteoporosis as defined by a *T* score for BMD of less than or equal to −2.5 SD, though patients with fragility fractures could receive treatment in the absence of BMD. The approach was highly specific, but lacked sensitivity since most fractures would occur in individuals with a BMD that exceeded a *T* score of −2.5 SD [28–31]. Moreover, the algorithm did not take account of the fact that some of the risk indicators are associated with an additional fracture risk over and above that captured by BMD.

FRAX®

Since the development of these guidelines, it has been recognised that the combination of information from independent risk factors for fracture improves the ability to characterise risk [32–34]. Risk factors for fracture that contribute independently of BMD include age, sex, a prior fragility fracture and a range of clinical risk factors. More recently, the independent contribution of different risk factors for fracture has been quantified [24, 35] permitting the calculation of absolute risk with the FRAX® tool (http://www.shef.ac.uk/FRAX).

FRAX uses easily obtained CRFs to estimate 10-year fracture probability. The estimate can be used alone or with femoral neck BMD to enhance fracture risk prediction. In addition, FRAX uses Poisson regression to derive hazard functions of death as well as fracture. These hazard functions are continuous as a function of time which permits the calculation of the 10-year probability of a major osteoporotic fracture (hip, clinical spine, humerus or wrist fracture) and the 10-year probability of hip fracture. Some of the risk factors affect the risk of death as well as the fracture risk. Examples include increasing age, low body mass index (BMI), low BMD and smoking. Most other risk engines calculate the probability of a clinical event (e.g. a myocardial infarct) without taking into account the possibility of death from other causes. In addition, the FRAX model has been calibrated for different countries [24, 35, 36].

Probability of fracture is calculated in men or women from age, BMI computed from height and weight, and dichotomised risk variables that comprise;

A prior fragility fracture
Parental history of hip fracture
Current tobacco smoking
Ever long-term use of oral glucocorticoids
Rheumatoid arthritis
Other causes of secondary osteoporosis
Daily alcohol consumption of 3 or more units daily

Femoral neck BMD can additionally be entered as a machine-specific BMD or as a *T* score derived from the NHANES III database for female Caucasians aged 20–29 years [37]. When entered, calculations give the 10-year probabilities as defined above with the inclusion of BMD.

The integration of risk factors increases the sensitivity of assessment (i.e. detects more patients who will fracture) without sacrificing specificity [38–40].

National Osteoporosis Guideline Group

For the reasons above, the National Osteoporosis Guideline Group (NOGG) in the UK published guidelines in 2008 recommending that decisions about treatment be based on the probability of fracture and not on the *T* score (http://www.shef.ac.uk/NOGG/index.htm) [41, 42]. The risk assessment is initially based on age and risk factors alone. Individuals with one or more of the risk factors used in the FRAX algorithm or a low BMI (≤19 kg/m^2) are eligible for the assessment of fracture probability. In some individuals close to a threshold value, a measurement of BMD is recommended to refine the estimate of risk [35, 42]. The intervention threshold, expressed as the 10-year probability of a major osteoporotic fracture ranges from 7.5% at the age of 50 years to 30% at the age of 80 years (Fig. 1). This threshold, equivalent to the fracture probability of a woman with a prior fragility fracture, was modelled on the RCP guidelines that recommended treatment in women with a prior fragility fracture.

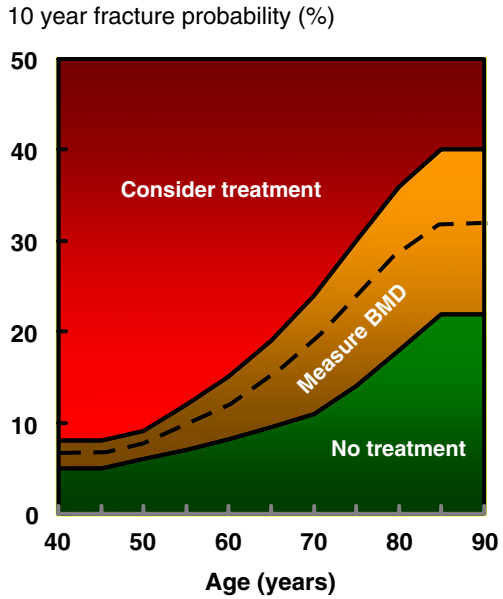

Fig. 1 Assessment guidelines of the National Osteoporosis Guideline Group based on the 10-year probability of a major fracture (%). The dotted line denotes the intervention threshold. Where assessment is made in the absence of BMD, a BMD test is recommended for individuals where the probability assessment lies in the orange region (adapted from [42])

The NOGG guideline has been validated by health economic analysis based on the cost of generic alendronate which is considered to be first-line treatment [15]. The cost-effectiveness of alendronate directed to women at several intervention thresholds is shown in Table 1. The cost of medication was assumed to be £95 per annum (as given in the British National Formulary at the time of analysis, but now is about a quarter of this cost).

In women with osteoporosis (i.e. a femoral neck *T* score equal to −2.5 SD), the incremental cost-effectiveness ratio (ICER) was stable up to the age of 60 years and, thereafter, decreased progressively with increasing age. Treatment was cost-effective at all ages, even assuming a willingness to pay (WTP) of £20,000/quality-adjusted life years (QALY). Treatment was also cost-effective at all ages in women who had previously sustained a fragility fracture with a BMD set at the threshold of osteoporosis. Indeed, treatment was cost saving from the age of 75 years. A prior fragility fracture was a sufficiently strong risk factor that treatment was cost-effective even in women without other risk factors in whom BMD was not known (see Table 1). The latter scenario provides a health economic assessment at the intervention threshold used by NOGG in Fig. 1.

The impact of the NOGG guideline has been compared with the previous guideline of the RCP. The NOGG strategy selects more women at younger ages and fewer women at older ages than the RCP guidelines (Table 2). Over all ages, NOGG identified a similar number of women at high risk compared with the RCP strategy (average 35.7% vs. 34.6% across all ages). The NOGG strategy required a lower number of scans at each age. For example, NOGG required only 3.5 scans at the age of 50 years to identify one case of hip fracture, whereas RCP required 13.9. At the age of 75 years, the corresponding numbers needed to scan were 0.9 and 1.5. The lower number of BMD tests meant that the acquisition costs for identifying a hip fracture case and the total costs (acquisition and treatment) per hip fracture averted were also lower [43].

Ibandronate, risedronate, raloxifene, strontium ranelate and zoledronic acid are considered second-line treatments to be used in patients who cannot take or tolerate alendronate. Teriparatide is reserved for a small number of patients with multiple previous fractures.

The second-line treatments are more costly than generic alendronate and are, not surprisingly, less cost-effective (Table 3). It can be argued, therefore, that there are some patients who cannot take alendronate that are at a too-low risk to start a second-line treatment purely based on cost-effectiveness. This sets an ethical dilemma for the primary care physician in that patients who cannot take alendronate would not be afforded any treatment until their condition had deteriorated sufficiently to provide an alternative

Table 1 Cost-effectiveness of intervention with alendronate in women at the threshold of osteoporosis, with or without a prior fracture and in women with a previous fracture without BMD [15]

Age (years)	Cost (£000)/QALY gained		
	T score=−2.5 no previous fracture	*T* score=−2.5+ previous fracture	No BMD+previous fracture
50	14.7	6.7	14.6
55	16.2	7.3	14.1
60	14.3	7.3	11.6
65	7.0	2.9	5.0
70	3.7	0.8	2.1
75	3.0.	c.s.	c.s.
80	c.s.	c.s.	c.s.

BMI set to 26 kg/m^2

c.s. cost saving

treatment. In order to avoid the problem, NOGG used the same intervention thresholds for these second-line agents as used for generic alendronate despite their higher price. This position is taken because cost-effective scenarios for these second-line interventions are found given a WTP of £20,000–30,000/QALY (see Table 3) [15, 18–20, 44–46].

It is of interest that the cost-effectiveness of alendronate in the scenarios above was modelled using a cost of £95 per year. By the time that the analysis was published, the cost had decreased to £45 per year and is now £26 yearly. Lower price is an argument for extending treatment to patients at lower risk. Total costs may increase or decrease. Although this is not relevant from a cost-effectiveness perspective, it may give headroom for innovation; i.e. resources available for new treatments (within or outside osteoporosis). If resources were to be allocated to osteoporosis, then the headroom is considerable. Assume, for example, that all treatments have equal efficacy if used in the correct population, and assume that the cost of second-line treatment is £300 per year (not teriparatide or human recombinant PTH (1–84)), then the proportion of patients (W) that could be offered innovative treatment can be quantified as;

$$W \cdot 300 + (1 - W) \cdot 45 = 95$$
$$W = 0.196$$

Or with a price of alendronate set at £26/year

$$W \cdot 300 + (1 - W) \cdot 26 = 95$$
$$W = 0.252$$

Thus, up to 19.6% or 25.2% of patients could be receiving other treatments (not PTH peptides) afforded by this headroom [42]. A very similar conclusion is obtained when a health economic model is used [47].

In reality, the use of agents other than alendronate in the UK is confined to a minority and continues to decrease as a proportion of market share [48]. In 1998, alendronate accounted for 14% of the prescription market which has increased progressively thereafter to 70% 10 years later (Fig. 2). The progressive rise, unaffected by generics (introduced in 2005) or NICE appraisals (2005), suggests that the practice of primary care physicians is likely to fall within the demands for cost-effectiveness.

NICE appraisal

The approach used by NICE differs in several fundamental ways from existing guidelines provided by NOGG [12, 13]. First, the remit of its guidance applies only to postmenopausal women with a *T* score of less than or equal to −2.5 SD, does not consider women with low bone mass but with other strong risk factors for fracture, does not include men and does not provide for glucocorticoid-induced osteoporosis. Second, recommendations are not available for zoledronic acid and ibandronate which have not been formally appraised. Third, the guidance makes a distinction between primary and secondary prevention of fractures, valuing secondary prevention higher than primary prevention in terms of WTP. Fourth, the guidance gives intervention thresholds on the basis of a *T* score for BMD

Table 2 Comparison of the total costs (identification and treatment) per hip fracture averted and the number of high risk women identified with each of the strategies [43]

Age (years)	Cost per hip fracture averted (£)		Number selected/1,000	
	RCP	NOGG	RCP	NOGG
50	6,210	4,797	10	22
55	4,607	3,678	18	16
60	3,504	3,020	21	14
65	2,255	2,144	25	38
70	1,764	1,716	35	29
75	1,609	1,537	45	18
80	1,371	1,306	49	15
85	1,306	1,231	56	15

Table 3 Analysis of the cost-effectiveness of interventions in women aged 70 years from the UK [15]

Intervention	Cost (£)/QALY gained		
	T score=−2.5 no previous fracture	*T* score=−2.5+ previous fracture	No BMD[a]+previous fracture
Alendronate	3,714	867	2,119
Etidronate	12,869	10,098	9,093
Ibandronate daily	20,956	14,617	14,694
Ibandronate intermittent	31,154	21,587	21,745
Raloxifene	11,184	10,379	10,808
Raloxifene without breast cancer	34,011	23,544	23,755
Risedronate	18,271	12,659	13,853
Strontium ranelate	25,677	18,332	19,221
Strontium ranelate, post hoc analysis	18,628	13,077	13,673

[a] Conservatively, all patients treated were assumed to have a BMD test before treatment and 2-yearly thereafter

rather than on fracture probability. As reviewed below, this is problematic because multiple intervention thresholds are provided depending on the agent that it is intended to use, the presence or absence of a prior fragility fracture, age, and the number and type of additional clinical risk factors present.

In common with the NOGG guidance, there is an overlap in CRFs considered by NICE and the FRAX risk factors, but NICE categorises them in the context of guidance. The risk factors are categorised by NICE as risk factors for fracture (category A, Table 4) and risk factors for low BMD (category B). The many other causes of secondary osteoporosis are not considered.

Both NOGG and NICE guidance recommend alendronate as a first-line treatment. Intervention thresholds recommended by NICE for alendronate are given in Tables 5 and 6. Two tables are provided because the intervention thresholds differ between primary and secondary prevention. For primary prevention, alendronate is recommended as a treatment option in postmenopausal women younger than 65 years where they have a CRF for fracture (category A, Table 4) and at least one CRF for low BMD (category B, Table 4) and who are confirmed to have osteoporosis (that is, a *T* score of −2.5 SD or lower). For women aged 65–69 years, a Category A CRF is required and a *T* score of −2.5 SD or lower. Women aged 70 years or older must have a CRF (category A or B) and a *T* score of −2.5 SD or lower. Thus, the intervention threshold is a *T* score of −2.5 SD together with specific requirements for CRFs. An exception is that provision is made for women aged 75 years or older who have two or more independent CRFs (category A or B), in whom a BMD test may not be required if the responsible clinician considers it to be clinically inappropriate or unfeasible.

In those women who cannot take alendronate, the use of alternative options (etidronate, risedronate, or strontium ranelate) requires more stringent *T* score criteria, and these change with age. For example, a woman aged 65–69 years with a category A CRF must have a *T* score of −3.5 SD or lower to qualify for etidronate or risedronate and a *T* score of −4.0 SD or lower to be allowed strontium ranelate.

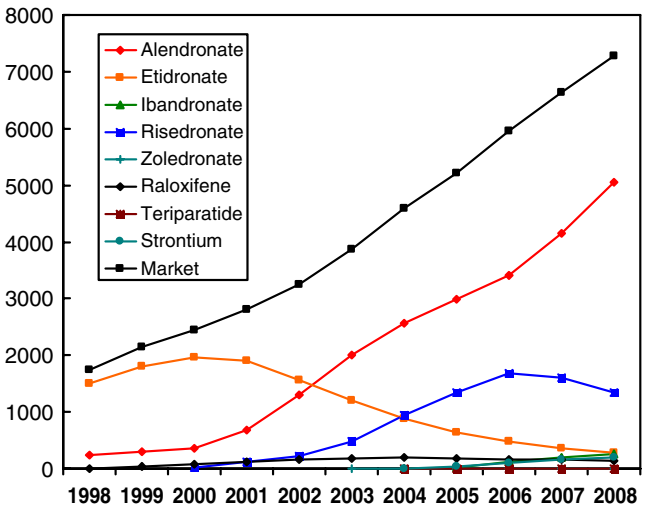

Fig. 2 Prescribing volume for osteoporosis treatments in England 1998–2008 [48]

Table 4 Risk factors for fracture (A) or low BMD (B) used in the NICE guidance [12, 13]

A Fracture	Parental history of hip fracture
	Alcohol intake of 4 or more units per day
	Rheumatoid arthritis
B Low BMD	Low body mass index (<22 kg/m^2)
	Ankylosing spondylitis
	Crohn's disease
	Prolonged immobility
	Untreated premature menopause
	Rheumatoid arthritis

Table 5 Treatment recommendations for primary prevention of fractures by treatment in postmenopausal women according to BMD *T* score and the number and type of clinical risk factors (CRFs)

		Number of clinical risk factors					
		0		1		2	
Agent	Age	CRF	BMD	CRF	BMD	CRF	BMD
Alendronate	<65	–	–[b]	–	–[b]	A and B	≤−2.5
	65–69	–	–[b]	A only	≤−2.5	A only	≤−2.5
	70+	–	–[b]	A or B	≤−2.5	A or B	≤−2.5
	75+[a]	–	–[b]			A or B	–
Risedronate[c]	<65	–	–[b]	–	–[b]	–	–[b]
	65–69	–	–[b]	A only	−3.5	A only	−3.0
	70–74		−3.5	A only	−3.0	A only	−2.5
	75+		−3.0	A only	−3.0	A only	−2.5
	75+*					A or B	–
Etidronate[c]	<65	–	–[b]	–	–[b]	–[b]	–
	65–69	–	–[b]	A only	−3.5	A only	−3.0
	70–74		−3.5	A only	−3.0	A only	−2.5
	75+		−3.0	A only	−3.0	A only	−2.5
	75+*					A or B	–
Strontium ranelate[d]	<65	–	–[b]	–	–[b]	–	–[b]
	65–69	–	–[b]	1	−4.5	A only	−4.0
	70–74		−4.5	1	−4.0	A only	−3.5
	75+		−4.0	1	−4.0	A only	−3.5
Raloxifene	<65	–	–[b]	–	–[b]	–	–[b]
	65–69	–	–[b]	–	–[b]	–	–[b]
	70–74	–	–[b]	–	–[b]	–	–[b]
Teriparatide	<65	–	–[b]	–	–[b]	–	–[b]
	65–69	–	–[b]	–	–[b]	–	–[b]
	70–74	–	–[b]	–	–[b]	–	–[b]
Ibandronate			No recommendations—not appraised				
Zoledronic acid			No recommendations—not appraised				

For the relevant CRFs to take into account (A and/or B), see Table 4 (extracted from [12]).

[a] In women aged 75 years or older who have two or more independent clinical risk factors for fracture or indicators of low BMD, a DXA scan may not be required if the responsible clinician considers it to be clinically inappropriate or unfeasible

[b] Treatment is not recommended irrespective of BMD

[c] Risedronate and etidronate are recommended in postmenopausal women who are unable to comply with the special instructions for the administration of alendronate, or have a contraindication to or are intolerant of alendronate

[d] Strontium ranelate is recommended in postmenopausal women who are unable to comply with the special instructions for the administration of alendronate and either risedronate or etidronate, or have a contraindication to or are intolerant of alendronate and either risedronate or etidronate

Neither raloxifene nor teriparatide are recommended by NICE for primary prevention.

For secondary prevention, all postmenopausal women with a prior fragility fracture may receive alendronate if they are aged over 75 years; younger postmenopausal women with a fracture must be shown to have a *T* score less than or equal to −2.5 SD before being eligible for treatment (Table 6), regardless of whether or not CRFs are present.

In those women who cannot take alendronate, the use of alternative options (etidronate, risedronate, strontium ranelate or raloxifene) requires more stringent criteria that also vary with age. For example, a woman aged 65–69 years with a fragility fracture must have a *T* score of −3 SD or lower to qualify for etidronate or risedronate and a *T* score below −4.0 SD to be allowed strontium ranelate or raloxifene. Teriparatide is reserved for women with multiple fractures and a *T* score threshold that ranges from −4.5 to −5.0 SD. Etidronate is positioned alongside risedronate as a second-line option, even though evidence for its efficacy to reduce spine fractures would not be regarded as

Table 6 Treatment recommendations for secondary prevention of fractures in postmenopausal women by treatment according to BMD and clinical risk factors

		Number of clinical risk factors					
		0		1		2	
Agent	Age	CRF	BMD	CRF	BMD	CRF	BMD
Alendronate	50+		≤−2.5				
	75+[a]	–	–				
Risedronate[c]	50–54	–	–[b]	A only	−3.0	A only	−2.5
	55–59	–	−3.0	A only	−3.0	A only	−2.5
	60–64		−3.0	A only	−3.0	A only	−2.5
	65–69		−3.0	A only	−2.5	A only	−2.5
	70+		−2.5	A only	−2.5	A or B	−2.5
Etidronate[c]	50–54	–	–[b]	A only	−3.0	A only	−2.5
	55–59	–	−3.0	A only	−3.0	A only	−2.5
	60–64		−3.0	A only	−3.0	A only	−2.5
	65–69		−3.0	A only	−3.0	A only	−2.5
	70+		−2.5	A only	−2.5	A or B	−2.5
Strontium ranelate[d]	50–54	–	–[b]	A only	−3.5	A only	−3.5
	55–59		−4.0	A only	−3.5	A only	−3.5
	60–64		−4.0	A only	−3.5	A only	−3.5
	65–69		−4.0	A only	−3.5	A only	−3.0
	70–74		−3.0	A only	−3.0	A only	−2.5
	75+		−3.0	A only	−2.5	A only	−2.5
Raloxifene[d]	50–54	–	–[b]	A only	−3.5	A only	−3.5
	55–59		−4.0	A only	−3.5	A only	−3.5
	60–64		−4.0	A only	−3.5	A only	−3.5
	65–69		−4.0	A only	−3.5	A only	−3.0
	70–74		−3.0	A only	−3.0	A only	−2.5
	75+		−3.0	A only	−2.5	A only	−2.5
Teriparatide[e]	50–55	–	–[b]				
	55–64		−4.0[f]				
	65+		−4.0				
	65+		−3.5[f]				
Ibandronate			No recommendations—not appraised				
Zoledronic acid			No recommendations—not appraised				

For the relevant clinical risk factors (CRF) to take into account (A and/or B), see Table 4 (extracted from [13])

[a] In women aged 75 years or older, a DXA scan may not be required if the responsible clinician considers it to be clinically inappropriate or unfeasible.

[b] Treatment is not recommended irrespective of BMD.

[c] Risedronate and etidronate are recommended in postmenopausal women who are unable to comply with the special instructions for the administration of alendronate, or have a contraindication to or are intolerant of alendronate

[d] Strontium ranelate or raloxifene are recommended in postmenopausal women who are unable to comply with the special instructions for the administration of alendronate and either risedronate or etidronate, or have a contraindication to or are intolerant of alendronate and either risedronate or etidronate

[e] Teriparatide is recommended as an alternative treatment who are unable to take alendronate and either risedronate or etidronate, or have a contraindication to or are intolerant of alendronate and either risedronate or etidronate, or who have a contraindication to, or are intolerant of strontium ranelate, or who have had an unsatisfactory response to treatment with alendronate, risedronate or etidronate

[f] Three or more prior fractures plus BMD requirement

[g] Not clear whether CRF required

robust by modern standards and prospective evidence for non-vertebral fracture reduction is nonexistent.

The comparison of alendronate with other treatments is fraught with difficulties. A major problem is that efficacy of each agent is taken from meta-analyses of randomised controlled trials (RCTs) in the absence of comparator studies that evaluate fractures as the primary outcome. The baseline characteristics, including fracture risk vary widely between studies and there is reason to suppose that responsiveness to an intervention differs according to the type of patient enrolled. Examples are provided in the Fracture Intervention Trials with alendronate [49, 50] and the hip fracture studies with risedronate [51] where the relative risk reductions varied between active treatment arms. More recently, a greater efficacy of clodronate and bazedoxifene has been observed in patients with the higher pre-treatment fracture probabilities as assessed by FRAX [52, 53]. These considerations suggest that in the context of cost-effectiveness, there is much greater uncertainty over incremental efficacy than incremental costs or even comparative persistence. With these limitations, it is reasonable that alendronate is considered a first-line treatment, but that second-line agents may be a cost-effective alternative in patients that are at higher risk of fracture or likely to discontinue treatment.

These considerations apart, the approach by NICE to target second-line treatments is inappropriate in the setting of patients that cannot tolerate alendronate or in whom the drug is contraindicated. Alendronate is not a relevant comparator in the cost-effectiveness analysis if it cannot be used. A comparison with no treatment is the relevant comparison for patients that will not adhere to oral therapy and stop treatment after a rather short time.

As previously noted, the use of different *T* score thresholds raises difficult practical issues for physicians and particularly discriminates against the frail and elderly, who are most likely to be intolerant of alendronate. Thus, women who are unable to tolerate alendronate must wait until their BMD has decreased or they have had a fracture before they can receive an alternative treatment (costing less than £350/year). Furthermore, the complexity of the criteria required for second-line options in individual women makes it unworkable in the primary care setting. Ultimately, the GDG of NICE has the task of translating the appraisals into guidelines. At the time of writing, the GDG is suspended but the complexity of the appraisal by NICE will not make this an easy task if it reconvenes.

The NICE model

The economic model used by NICE was based in Excel. The structure, data and assumptions used have been described in health technology assessment (HTA) reports [54, 55]. The model estimates the cost-effectiveness based on Gaussian regression functions that are derived from an individual state transition model. The Gaussian functions were estimated by simulating the cost-effectiveness over intervals for several of the input parameters in the individual state transition model [56]. The Gaussian functions cannot be evaluated since the individual state transition model was not provided by NICE for consultation. Thus, it has not been possible to evaluate the model fully and it cannot be considered, therefore, to be fully executable (following a request to NICE, the individual state transition model was forwarded to consultees, but no extra time was afforded to evaluate it and comments on its functionality embargoed). In addition, the opinion of the NICE was that the individual state transition model was not the relevant model because all the outputs could be derived from the Gaussian functions. This appears to be an extraordinary position given the great detail afforded in the HTA report to the individual state transition model [55]. The position implies that the two models are to be viewed as separate entities and have no direct connection. In reality, they are very much linked since the ghost model relies on the Gaussian functions that were estimated using the previous model. The current model would be incapable of producing any ICERs without the existence of the previous model.

The reliance on a 'ghost model' does not make it possible to question the individual simulation model or any of the data used in its construction, even though this would be a critical step in verifying its adequacy and accuracy. Thus the "model" supplied does not fit with any description of a "fully executable" unedited model even if these outputs were the only ones used to guide the appraisal.

The use of Gaussian regression functions also gives rise to inflexibility and several variables cannot be changed to undertake sensitivity analyses. These comprise:

- Discount rate for QALYs
- Discount rate for costs
- Body mass index
- Mortality adjustments to the general population
- Mortality adjustments in the presence of clinical risk factors
- Baseline population risk of fracture
- Time horizon
- Combinations of CRFs other than 19 pre-specified combinations

The NICE appraisal stated that these were intentionally fixed [12, 13]. What could be reviewed were those components that had been added on top of the Gaussian functions and which were not included in the individual state transition model. The following variables could be changed

- Efficacy of treatment
- Efficacy related to additional CRFs
- Drug costs
- Fracture-related costs
- Disutility associated with fractures
- Annual fracture risk at start of treatment
- Compliance (proportion of patients that stop within the first 6 months)

The process for obtaining the ICER from the individual simulations to the final ICER in the NICE model is depicted in Fig. 3.

The final ICER, as described in the figure above, for a defined patient group was not the end output for the interpretation of the cost-effectiveness of a treatment. After all the ICERs had been estimated for different numbers of CRFs, they were grouped together and the costs related to strategies for identifying these patients (primarily in the assessment of prevention treatment) were added (details reviewed later).

External review of the ghost model was undertaken by four consultees, but under a confidentiality clause that prohibited release of data generated by the consultees. Notwithstanding some general observations can be made from the appraisals and observations previously available in the public domain.

Validation of the model

When developing cost-effectiveness models it is important to validate the model to ensure that it is both internally and externally rigid. Internal validation is required to ensure that the model calculates correctly according to its specification and the data used. External validation ensures that the model accurately reflects epidemiology (e.g. fracture risk and mortality), treatment effect and characteristics of the target patient groups both in terms of the data, assumptions and model structure. Internal validation can be conducted in several ways. One method is to rebuild the model in another software (sometimes also done by a different programmer) in order to replicate results to ensure that there are no programming errors. Another approach is to compare the outputs (such as fracture risks) of the model to other estimations of the output using the same or similar data.

External validation is also sometimes referred to as methodological uncertainty, which arises when comparing study results based on different methods. This most often originates in a disagreement between researchers about the most appropriate method, data and assumptions to be used. This type of uncertainty is often best handled by sensitivity analysis and agreement upon a reference case model.

Unfortunately, the NICE model provided did not have the simulated risk as an output of the model so that it is not possible to validate the model through estimated fracture risks or mortality. Nor is it possible to determine the accuracy with which the model reproduced the epidemiology of osteoporosis in the UK. NICE note [12, 13] that an HTA report and a further paper [55, 57] were peer reviewed, but neither of these papers assessed the validity of the model. Nevertheless, NICE declared itself satisfied with the validity of the model since differences between the

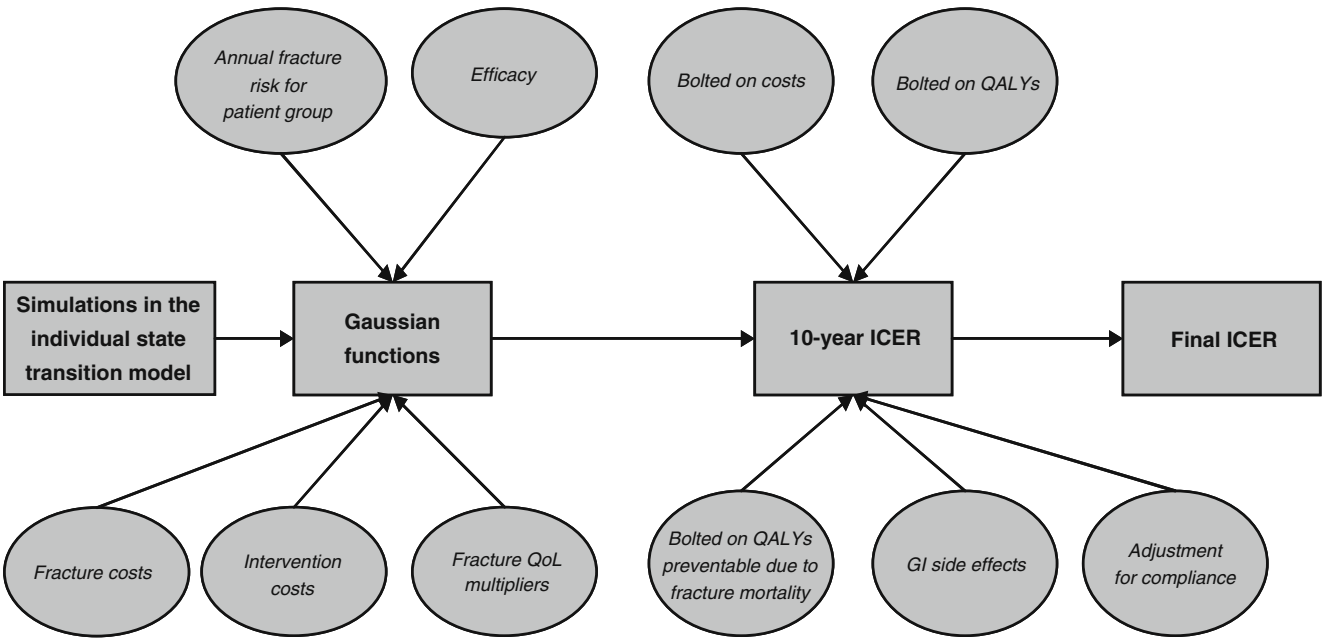

Fig. 3 The process to obtain the ICER in the NICE model

results obtained using an alternative model and the Assessment Group's model were largely because of differences in the assumptions used [12, 13]. The evidence available challenges this view as detailed below.

Comparison of models

The Assessment Group for the NICE appraisals believes that the validity of the model structure can be inferred by comparison with another published osteoporosis model that has been used as a reference model for the International Osteoporosis Foundation [58]. They note that the results produced by the 'NICE' model and the reference model are similar when populated with similar input parameters with regard to the cost-effectiveness of alendronate (witness statement of M Stevenson to the High Court, January 2009). The witness statement also notes that the adaptations made to the model to allow for effects beyond the initial 10-year time horizon appear to be appropriate. These data, in the form of a letter to Osteoporosis International were not accepted for publication.

Unfortunately, the argument fails on several counts. The first is that this is not a test of external validity. The second is that the results presented to the court do not show concordance (Table 7). It is notable that the conclusion is based on only 11 numerical examples. Also, the sampling frame is biassed by only considering very cost-effective scenarios. It cannot draw any inference that cost-ineffective scenarios using the NICE model would also be cost-ineffective using the reference model. In any event, there appear to be large numerical discrepancies when comparing models. For, these reasons, it is not possible to support the conclusions of the Assessment Group, and the inadequacy of the argument may be the reason why the letter was not accepted for publication.

Time horizon used by NICE

Several health economic assessments have drawn attention to discrepancies in estimates of cost-effectiveness produced by NICE and other models [15–17, 19, 20, 59]. It is difficult to determine why the results differ, but ultimately reasons reside in either the construct of the model or the assumptions used to populate the model. With regard to construct, the NICE model uses predominantly a 10-year time horizon rather than considering the lifetime of the patient as is the norm for chronic diseases. The use of a 10-year horizon has a large effect on apparent cost-effectiveness [15, 59] since, for example, an individual who dies after 9 years is dead for life, and not for 1 year, as would be assumed with a 10-year horizon.

In order to overcome this deficit, the NICE model preserved the time frame but 'bolted on' adjustments to overcome this flaw in the model construct. Two types of bolt-on factors were used in the model to adjust the incremental values estimated from the Gaussian functions. The first bolt-on adjusted treatment-related decreases in mortality to extend beyond the 10-year time horizon. The second adjustment was related to additional QALYs gained beyond 10-years to account for preventable deaths due to avoided fractures during the treatment period of 5 years [55]. In short, the expected remaining QALYs for a patient alive at the end of

Table 7 The cost per QALY gained (£000) for women treated with alendronate at a *T* score of −2.5 SD with and without a previous fracture from a previously published analysis [15] and the NICE model populated with the same assumptions (witness statement of M Stevenson to the High Court, London; January 2009)

Age (years)	Kanis et al. [15]	NICE model	Difference (%)
T score=−2.5 SD no previous fracture			
50	14.7	26.0	+77
55	16.2	21.0	+30
60	14.3	17.7	+24
65	7.0	14.0	+100
70	3.7	6.1	+65
75	3.0	1.7	−43
T score=−2.5 and previous fracture			
50	6.7	8.5	+27
55	7.3	7.4	+1
60	7.3	6.6	–10
65	2.9	5.0	+72
70	0.8	1.4	+75
75	c.s	c.s	–

c.s. cost saving

10-years were multiplied with the number of potentially prevented fracture deaths during the 5-year treatment period.

These adjustments only related to preventable deaths during the 5 years of treatment. However, during the offset period after the intervention, where a residual effect of treatment is assumed, there should be an impact on the number of preventable deaths which may not have been accounted for in the NICE model. Also, the number of fractures and deaths will differ between the comparator interventions even after the 10 years which have an impact on both QALYs and costs which seem not to be accounted for in the model.

However, in the model, there were several additional bolt-ons, two of which (*wristbonusat2.5* and *phbonusat2.5*) were mentioned in the appraisals, but the functionality of which were neither described in any published report nor in the appraisals [12, 13, 55].

Unfortunately, there are no data available that test the sensitivity of the NICE model to changes in the time horizon and no way to test the adequacy of the bolt-on to overcome the intrinsic deficit in the model.

Adjustments for compliance

In the HTA reports [54, 55], it is assumed that 50% of patients stop treatment within the first month. The patients that drop out of treatment are not simulated in the model. The patients that are simulated in the model are only those that persist on treatment for the whole intervention period. This is probably because compliance functionality was not implemented at the time it was decided to produce the Gaussian functions. Instead, an adjustment was made on the cost side to account for non-compliers by adding on one additional month of intervention costs. Any adjustment on the effect side is not necessary since non-compliers were not assumed to have any effect of treatment. A problem with such an approach is that those who discontinue treatment are likely to do so at time points throughout the 5-year period and should thus receive some health benefit, as well as additional drug costs. Patients who persist longer will have the benefit of a longer offset time. The NICE approach to account for compliance will overestimate both the incremental costs and QALYs gained [60] so that there may not be a major impact on the ICER compared to an approach where all patients are simulated in the model. This has not, however, been tested by NICE.

The use of risk factors in the NICE model

Annual risk of fracture

The annual risk of fracture was computed by NICE from the data supplied by the World Health Organization (WHO) Collaborating Centre for Metabolic Bone Disease at the University of Sheffield. The FRAX® algorithm uses fracture hazards and death hazards to compute 10-year fracture probabilities for any combination of clinical risk factors (CRFs). Regrettably, the NICE model does not permit the calculation of 10-year fracture probabilities, despite advice from the GDG and other consultees to the contrary, so that the integrity of the NICE application of FRAX® cannot be directly addressed.

By contrast, the NICE model used a 1 year time frame. The annual risks were entered directly as values in the excel sheets and it is not possible, therefore, to evaluate how the actual calculation of the risks was derived. NICE noted that there were discrepant values for fracture probability as calculated by NICE and by the consultees [12, 13] but the decision support unit (DSU) suggested that the differences in the estimates of fracture risk obtained using the FRAX fracture risk calculation tool and the Assessment model did not necessarily suggest that the WHO algorithm had been incorrectly applied.

Mortality

The FRAX® algorithms can also be used to assess the probability of death related to any combination of CRFs, i.e. FRAX® can be used to adjust the mortality for a specific patient group. This part of FRAX® was not implemented in the NICE model [55]. NICE confirmed that increases in mortality associated with clinical risk factors were not taken account of in the model in the interests of simplicity [12, 13]. The inclusion of mortality effects increase the ICERs for women with clinical risk factors because fewer QALY benefits would accrue in the model for women who die of causes related to risk factors. Conversely, patients with CRFs that are not associated with excess mortality would accrue more QALYs. Indeed, survival is significantly higher the fewer the CRFs [15]. The NICE appraisals state ‘the overall effect of including the increased mortality associated with clinical risk factors would be small’ but, even if true, misses the point that the error of accuracy will deny some patients the benefit of treatment.

Body mass index

NICE used a fixed BMI in the computations of fracture risk [55] set at 26 kg/m^2 for all simulations and could not be changed. BMI was also used as a dichotomous risk variable by NICE in their case-finding strategy. The threshold used was a BMI of 22 kg/m^2 [8, 9]. The effect of omitting BMI as a continuous variable on fracture probability is shown in Table 8 for women aged 70 years with a prior fracture. In the absence of BMD, the 10-year probability of a major fracture varied more than twofold,

Table 8 The effect of BMI on fracture probability for women aged 70 years with a prior fracture

BMI	*T* score −2.5		No BMD	
	Major	Hip	Major	Hip
15	16	4.7	26	14
20	19	5.3	22	8.4
25	22	5.8	20	4.8
30	21	5.4	17	3.8
35	20	5.0	15	2.9
40	19	4.6	13	2.3

Ten-year fracture probabilities are shown without including BMD and with a *T* score for femoral neck BMD is set at −2.5 SD. Data computed from the FRAX website (UK model, version 2.0)

ranging from 26% with a BMI of 15 kg/m^2 to 13% at a BMI of 40 kg/m^2. The range of hip fracture probabilities was even greater (from 2.3% to 14%). Variations were less marked, but still evident with the inclusion of BMD at a *T* score of −2.5 SD.

It is evident that the use of BMI as a fixed variable is not consistent with the construct of FRAX®. In addition, the deficit decreases the accuracy of all risk estimates except at the value used by NICE. The effect is very marked when BMD is not used to estimate risk. This has implications where management decisions are given for women without BMD (e.g. with a prior fracture aged 75 years or more). Though the impact is less, there are errors of accuracy incurred when BMD is added to the model. The error was acknowledged by the DSU which concluded, surprisingly, that a fixed BMI may favour treatment of women at risk of fracture compared with alternative BMI values. As in the case of mortality, the error of accuracy will deny some patients the benefit of treatment.

The use of a fixed BMI introduces other errors of accuracy in the computation of fracture probability. There is a significant interaction of BMI with BMI and for some outcomes with age [61]. In other words, the significance of a step change in BMI differs at different values of BMI and age. There is also a significant effect of BMI on mortality. The phenomenon is illustrated in Table 9 which gives the ratio of fracture probabilities at low values for BMI compared to average values (25 kg/m^2) at the ages of 50 and 70 years. At the age of 50 years and a BMI of 15 kg/m^2 the 10-year probability of a major fracture is increased by 40%. At the age of 70 years, the probability of a major fracture is decreased by 22%. These important interactions were not accommodated in the NICE model.

The potential impact of these omissions on cost-effectiveness is shown in Table 10 for a woman in a UK setting aged 70 years and a family history of hip fracture treated with strontium ranelate, using a model that incorporated the FRAX algorithms [19]. In the absence of BMD, cost-effectiveness ranged from £24,300 to £36,100/QALY gained over a modest range of BMI.

NICE argued that there is a significant but poor correlation between BMI and BMD and, for this reason, it was decided only to use BMD rather than BMI and BMD [12, 13]. This would only be a logical argument if there were a strong correlation, i.e. if BMI could be predicted from BMD which is clearly not the case.

Additional problems arise with the use of BMI in case finding (see *Cost-effectiveness of identification strategies*).

Intake of alcohol

The FRAX® model accommodates alcohol intake as a dichotomous risk variable. The threshold is set at an average intake of 3 or more units daily and is associated with an increased risk of hip fracture and a major osteoporotic fracture [62]. Notwithstanding, the NICE appraisal chose to use a threshold of 4 or more units daily. This is associated with a higher relative risk for fracture than the thresholds used by FRAX (Table 11). For example, the relative risk of hip fracture (with BMD) is 1.70 for an intake of >2 (i.e. 3 or more units) daily, but 2.05 at an average intake of 4 or more units daily. Thus, the use of the original FRAX® coefficient by NICE underestimates the fracture risk when the threshold is altered so that the cost-effectiveness of intervention will be underestimated.

Missing variables

The FRAX® model uses smoking and exposure to glucocorticoids as dichotomous risk variables. Neither risk factor was used by NICE. In the case of smoking, the reason cited was that the effect of smoking in women was not statistically significant when assessing risk of osteoporotic fractures taken as a whole [12, 13]. This view seems to be at variance with the published literature [63–65] including the data used to populate the FRAX model

Table 9 The effect of low BMI on fracture probability ratios for women aged 50 or 70 years with a prior fracture and with a *T* score for femoral neck BMD set at −2.5 SD

BMI	Age 50 years		Age 70 years	
	Major	Hip	Major	Hip
15	1.4	1.2	0.78	0.88
20	1.2	1.1	0.92	0.94
25	–	–	–	–

The ratio of 10-year fracture probabilities are shown at each BMI compared to a BMI of 25 kg/m^2 in an individual of the same age. Data computed from the FRAX website (UK model, version 2.0)

Table 10 The effect of low BMI on 10-year fracture probability and cost-effectiveness of strontium ranelate for women aged 70 years from the UK with a family history of hip fracture (model described in [19])

CRFs	BMI=20		BMI=26		BMI=32	
	Major	Hip	Major	Hip	Major	Hip
Absolute risk (%)	19.7	8.6	16.8	4.7	14.0	3.5
Relative risk	1.29	2.29	1.26	1.16	1.08	0.86
ICER (£000/QALY gained)	24.3	30.6	36.1			

(Table 12). Notwithstanding, NICE determined that the selection of risk factors to be used in the appraisal was a matter for NICE alone to consider and determine.

As was the case of smoking, exposure to glucocorticoids was not included as a risk factor in the assessments. The Committee did not consider it appropriate to include recommendations for women on long-term treatment with glucocorticoids because this group is at greatly increased risk of fracture, and therefore requires special consideration [12, 13]. The validity of the argument is questionable given that patients with a family history of hip fracture are a group at greatly increased risk of fracture, and therefore would also require special consideration (Table 13).

More likely, NICE could not be bothered to recognise the limitations of their model construct and any impact on cost-effectiveness scenarios. They couched this somewhat more elegantly in the appraisal as taking 'a pragmatic view that such amendments would have added unnecessarily to the mathematical complexity of an already complex clinical situation'. Even if these omissions had little effect overall on cost-effectiveness, they would have a marked impact on the accuracy with which patients close to an intervention threshold are considered eligible or ineligible for treatment.

Use of risk factors to compute ICERs

The ICER, as described in Fig. 3 above, for a defined patient group was not the final output for the interpretation of the cost-effectiveness of treatment. After all ICERs had been estimated for different numbers of CRFs, they were grouped together. Thus, whereas FRAX® provided the mechanism to compute the cost-effectiveness according to the specific risk factor, NICE weighted all risk factors equally.

The impact of this on fracture probability is shown in Table 14. For example, the average 10 year probability for women aged 65 years with two risk factors and a *T* score of −2.0 SD is 20%, but varies more than twofold (13% to 29%) depending on the risk factors.

A similar situation pertains when CRFs are accorded equal weights in the absence of BMD. For example, the average 10 year probability for women aged 65 years with two risk factors and a BMI of 25 kg/m^2 is 19%, but varies more than twofold (11% to 29%) depending on the risk factor. Other examples are given in Table 15 and on the FRAX® web site.

A similar inaccuracy results from the presentation of age and BMD in categories. Thus, NICE presented ICERs in age bands (e.g. 55–59 years) and *T* score bands (e.g. *T*= −3.0 to −3.5 SD). This makes direct comparisons with the results of NICE problematic because a mean value will differ from a point estimate at a specific age and a specific BMD.

For example, cost-effectiveness for strontium ranelate was given at £57,500/QALY for women with a prior fracture aged 55–59 years, with a *T* score that ranged between −3.0 and −3.5 SD and no clinical risk factors [55]. In the presence of one additional clinical risk factor (assumed to be a prior fracture in the context of the NICE appraisal), the cost-

Table 11 Risk ratio for fracture and 95% confidence intervals according to the intake of alcohol with and without adjustment for femoral neck BMD [62]

Consumption (units/day)	Without BMD		Adjusted for BMD	
	RR	95% CI	RR	95% CI
Osteoporotic fracture				
>2	1.38	1.16–1.65	1.36	1.13–1.63
>3	1.55	1.26–1.92	1.53	1.23–1.91
>4	1.70	1.30–2.22	1.64	1.24–2.27
Hip fracture				
>2	1.68	1.19–2.36	1.70	1.20–2.42
>3	1.92	1.28–2.88	2.05	1.35–3.11
>4	2.26	1.35–3.79	2.39	1.39–4.09

Table 12 Risk ratio for fracture (RR) and 95% confidence interval (CI) associated with current smoking by fracture outcome in men and women [64]

Outcome	Sex	RR	95%CI
Any kind of fracture	M	1.50	1.26–1.77
	F	1.18	1.07–1.30
	M+F	1.25	1.15–1.36
Osteoporotic fracture	M	1.53	1.27–1.83
	F	1.20	1.06–1.35
	M+F	1.29	1.17–1.43
Hip fracture	M	1.82	1.34–2.49
	F	1.85	1.46–2.34
	M+F	1.84	1.52–2.22

Table 13 The effect of selected clinical risk factors on fracture probability for women by age with a *T* score for femoral neck BMD set at −2.5 SD. BMI is set at 24 kg/m^2

Age	Risk factor			
	None	Prior fracture	Glucocorticoids	Family history
Major fracture				
50	6.2	12	10	11
60	9.4	16	15	17
70	14	21	21	22
80	16	23	24	32
Hip fracture				
50	1.6	3.3	2.9	1.7
60	2.4	4.3	4.4	2.5
70	3.7	5.7	6.6	7.8
80	6.1	8.0	10	24

Data computed from the FRAX website (UK model, version 2.0)

effectiveness ratio decreased to £46,800 and in the presence of two clinical risk factors was £34,000. The analysis gives an inaccurate estimate of cost-effectiveness, since it does not provide information at a specific *T* score (e.g. at −3.0 or at −3.5 SD) and a specific age (e.g. at 55 years or at 60 years). Moreover, the cost-effectiveness varies according to the specific risk factor whereas, as noted above, NICE weighted all risk factors equally.

The error of accuracy is illustrated from the example in Table 16 using a model that incorporated FRAX [19]. The cost-effectiveness of strontium ranelate for women with a prior fracture aged 55–59 years, with a *T* score that lay between −3.0 and −3.5 SD and two clinical risk factors was given by a single estimate in the NICE appraisal of £46,800/QALY gained [55]. In the FRAX-based model (Table 16), cost-effectiveness ranged from £19,200 to £30,100 depending on the *T* score, age and the nature of the clinical risk factor. In other words there was a greater than 1.5-fold variation in cost-effectiveness, covered by NICE as a single estimate.

Similar conclusions are reached using point estimates provided as an addendum to the FADs by NICE [66]. For example, the cost-effectiveness in a woman with a prior fracture and a *T* score of −3.0 SD at the age of 65 years was given as £38,499. With a 0.5 decrement in *T* score and 5 year increment in age, the ICER decreased to £14,986—a greater than twofold variation in cost-effectiveness.

NICE reports that the median ICER for a range of CRF combinations within an age span and *T* score interval was used for simplicity and 'the only practical way forward' to produce workable recommendations. NICE stated that the full FRAX model was unavailable at the time of the appraisal, but neglected to state that this was offered. The effect of using the median ICER is to favour those women who have a CRF which conferred a lower than median risk, but would disfavour women who have a CRF which conferred a higher than median risk. Thus NICE could not possibly deny that the "median" solution would differ to an ICER which is estimated by properly weighting the different CRFs. The same considerations apply to the *T* score, age and BMI. The fact that most of the errors are introduced after the model output and not by the model itself does not lessen the errors. Thus the decision of NICE to use age groups, median coefficients and *T* score groups decreases the accuracy of the information by which patients' risk and cost-effectiveness can be stratified.

The GDG and other consultees have consistently recommended that NICE report fracture probabilities and base intervention thresholds on probabilities using individual CRFs with their appropriate weightings. The argument that this is too complex is flawed and is negated by the development of the National Osteoporosis Guidelines by NOGG [23, 41] supported by many learned societies and patient support organisations such as the Royal College of Physicians and the National Osteoporosis Society. These guidelines provide practical advice based on the accurate assessment of fracture probability and are increasingly used throughout the UK [67]. Indeed the NOGG website (www.shef.ac.uk/NOGG) receives more than 11,000 hits daily.

Table 14 Ten-year probability of a major osteoporotic fracture (%) according to BMD *T* score at the femoral neck and the number of clinical risk factors (CRFs) in women aged 65 years from the UK [Data from FRAX® web site UK model, version 2.0]

Number of CRFs	BMD *T* score (femoral neck)					
	−4.0	−3.0	−2.0	−1.0	0	1.0
0	27	15	9.7	7.1	5.9	5.0
1	37 (33–41)	22 (18–26)	14 (10–18)	10 (7.1–14)	8.5 (5.7–12)	7.3 (4.8–10)
2	49 (42–58)	30 (23–40)	20 (13–29)	15 (8.6–23)	12 (6.8–19)	10 (5.6–17)
3	62 (53–72)	41 (30–55)	27 (17–42)	20 (11–34)	17 (8.7–29)	15 (7.2–26)
4	73 (63–81)	52 (42–65)	36 (26–51)	27 (18–41)	23 (14–36)	20 (11–32)

Values in brackets denote the range of probability depending on the weight of the risk factor

Table 15 Ten-year probability of osteoporotic fractures (%) according to body mass index (BMI) and the number of clinical risk factors (CRFs) in women aged 65 years from the UK [data from FRAX® web site UK model, version 2.0]

Number of CRFs	BMI (kg/m^2)						
	15	20	25	30	35	40	45
0	11	9.3	8.6	7.4	6.5	5.6	4.9
1	16 (12–21)	14 (10–18)	13 (9.2–16)	11 (7.9–14)	9.8 (6.9–12)	8.5 (5.9–11)	7.4 (5.1–9.5)
2	24 (16–34)	21 (13–31)	19 (11–29)	17 (9.8–26)	14 (8.4–23)	13 (7.3–20)	11 (6.3–18)
3	35 (24–49)	30 (19–45)	27 (16–43)	24 (14–38)	21 (12–34)	18 (10–30)	16 (8.7–27)
4	48 (35–62)	42 (30–57)	38 (26–54)	34 (22–49)	30 (19–44)	26 (16–39)	23 (14–35)

Values in brackets denote the range of probability depending on the weight of the risk factor

Probability-based guidelines have been or are being developed in many regions of the world [4, 41, 42, 68–81]. The international web site receives approximately 200,000 hits daily which suggests that many do not share the view of NICE that intervention thresholds are more appropriately based on *T* scores rather than on fracture probability.

A supplementary argument by NICE that it was not possible to develop intervention thresholds based on probabilities or that FRAX cannot be integrated into economic models is unfounded [19, 20, 82, 83]. Thus, NICE did not fully consider the manner by which the problem can be remedied (see "Intervention thresholds", below).

Cost-effectiveness of identification strategies

The estimated ICERs for specific combinations of CRFs were not directly used by NICE in the interpretation of the results. Rather, the cost-effectiveness of treatment included the costs of an identification strategy based on age, *T* score and number of CRFs. The first step in the evaluation of the cost-effectiveness of the identification strategy for a given age range was that the average incremental cost and QALYs gained for each *T* score range and combination of CRFs (grouped in 0, 1, 2 and 3 CRFs). To assess whether an overall identification strategy was cost-effective at a given age range, the incremental values were multiplied by the number in the population for England and Wales estimated to fall within each combination of *T* score level and number of CRFs at different ranges of age. The total costs for the identification strategy, which were derived by multiplying the costs for BMD measurement and physician time with population numbers, were then added to the total population incremental cost. The total identification costs for each number of CRFs were then summarised and divided by the total QALY gained to obtain the cost per QALY gained for the entire identification strategy at each age. If the cost per QALY gained was below the threshold value (£20,000) then the identification strategy was considered cost-effective.

There are several limitations in this approach. Firstly, an average ICER is used to determine the population that would be identified as suitable for treatment. The use of the average ICER assumes that the prevalence of each CRF is equal. This is clearly not the case as illustrated above [24, 42], and weighted averages should have been used.

A further error is that, in the derivations of the identification strategy, the NICE model also included the ICERs based on alcohol, smoking and exposure to glucocorticoids which were CRFs not considered to be relevant risk factors in the NICE appraisal. It further did not include a low BMI as a risk variable—a weakness acknowledged in the HTA report to disadvantage younger women with CRFs and a low BMI [55]. The error is acknowledged

Table 16 Estimates of cost-effectiveness for strontium ranelate in women aged 55–59 years with a *T* score of −3.0 to −3.5 SD according to the presence or absence of clinical risk factors (CRF) [15, 19 and unpublished data using the same model]

CRF	NICE	Present study			
	Age 55–59	Age 55		Age 60	
	T=−3.0 to −3.5	*T*=−3.0	*T*=−3.5	*T*=−3.0	*T*=−3.5
Base case					
Prior fracture	57.5	36.3	28.8	36.0	28.9
Additional CRF	46.8				
Prior fracture+alcohol	na	30.1	23.7	30.0	24.0
Prior fracture+parental history	na	22.3	19.2	22.2	19.2

na not available

by NICE, but the Appraisal Committee took the view that its correction would have added unnecessarily to the mathematical complexity of an already complex clinical situation.

A third error is that the distribution of clinical risk factors over *T* score and age assumed an identical prevalence of CRFs over the entire range of *T* score which is clearly inappropriate. Indeed women above a threshold of probability on the basis of CRFs have a *T* score that is approximately 1 SD lower than women below the threshold [84]. Thus, the distribution of risk factors by age did not conform to their known distribution [42, 85].

A further flaw is that the acquisition algorithm claimed to follow the guidance of the Royal College of Physicians. This guidance indicates that women with CRFs would be eligible for a BMD test, and treatment offered to those with a *T* score of −2.5 SD. But an important exception is given for women with a prior fragility fracture where intervention may be considered without recourse to BMD testing [25, 26]. The guidance of the RCP mirrors that of many other clinical guidelines in Europe and North America [4, 23, 72, 75, 76, 79, 86, 87]. The omission of this aspect of the guidance for women under the age of 75 years increases the requirement for BMD tests in the identification strategy and thus inflates the cost. For example, the number of BMD tests to identify a patient for treatment between the ages of 70–74 years is given as 4.6 with a WTP of £20,000 and 5.8 with a WTP of £30,000 [55, Table 59]. By contrast, when the approach used by NOGG that follows the RCP guidance is used for the same age range, the average requirement is 0.4 BMD scans per patient identified for treatment [42].

Conclusion

The NICE appraisal neglected the impact of CRFs on the death hazards which will give rise to discrepancies between estimates of fracture risk using FRAX® and the estimates derived by NICE. The NICE appraisal did not take account of all variable interactions intrinsic to FRAX®. The NICE model made inappropriate use of BMI, alcohol intake, age and *T* score for BMD that introduced errors of accuracy which impact significantly on the ICER. The NICE model neglected smoking and glucocorticoid exposure which adds to errors of accuracy. The NICE model made errors in calculating identification costs which were additionally inflated by departing from the guidance of the RCP.

Populating the NICE model

NICE rightly base their assessments of efficacy on a platform of evidence-based medicine. In the case of osteoporosis, the consistency with which this has been applied can be questioned.

Effectiveness

Where possible, NICE based estimates of efficacy on meta-analysis of randomised controlled trials. A small digression was the use of a single estimate of efficacy for alendronate and risedronate based on pooled data for these two drugs. In a much larger deviation from any evidence base, NICE assumed that intervention in women without clinical risk factors had greater efficacy than shown in clinical trials, and conversely assumed that intervention in women with clinical risk factors had lesser efficacy [8, 9]. Take, for example, a woman aged 65 years with a *T* score of −2.5 SD and no clinical risk factors for fracture. The probability of a major osteoporotic fracture is 12% with a body mass index of 23.8 kg/m^2 (www.shef.ac.uk/FRAX). For an intervention with an efficacy of say 50% (RRR=0.5), NICE would assume greater efficacy—say 55%. In the same woman who additionally had a family history, the fracture probability rises to 21%. The effects of treatment on the incremental risk (the difference between 12% and 21%) were assumed to be half that of the trial results (an efficacy of 25% in this example). The adjustment was set so that if intervention was

used in the phase III setting, the overall efficacy would remain unchanged. The manipulation needs to assume (unlikely, but untested) that the prevalence of clinical risk factors is the same in Phase III studies as in the general population. The justification is based on the view that treatment of women with CRFs is less effective at any given BMD. This has been shown to be untrue in the many phase III studies addressing this question (reviewed recently [35, 42]).

To test the hypothesis directly that a candidate risk factor identified a risk amenable to treatment, it would be necessary to recruit individuals selected on the basis of the risk factor(s) to an RCT. The risk factor that is best evaluated in this way is BMD, and indeed the majority of therapeutic studies have recruited on the basis of low BMD as recommended by regulatory agencies in Europe [88]. In recent years, other trials have recruited on the basis of age, gender, a prior vertebral fracture and current exposure to glucocorticoids irrespective of BMD, and have shown therapeutic effects similar to those noted in RCTs based on BMD selection [89–91].

For other individual risk factors, comparable data are lacking, but several considerations suggest that this concern is misplaced in the context of the FRAX® risk factors. First, several studies have shown that intervention in the general population induces therapeutic results similar to those expected in individuals selected to be at high risk [92–94]. Second, studies have shown no significant interaction between response to treatment and the presence or absence of the risk factors used in FRAX including age, height, family history of fracture, low body weight or BMI, smoking, alcohol intake or prior non-vertebral fracture [95–99]. Third, the clinical risk factors are not totally independent of BMD and when clinical risk factors alone are used in women aged 70 years or more, BMD is approximately 1 SD lower in the high risk group compared with a low risk group [84]. Perhaps the best evidence is that response to intervention in elderly women recruited from the general population is greater, the higher the probability of fracture estimated without the inclusion of BMD from FRAX® [52]. Similar findings are reported for the SERM bazedoxifene. In this phase III intervention study, relative risk reduction compared to placebo was greater in women with the higher baseline fracture probabilities [53]. These considerations suggest that the risk factors chosen are appropriate in that they identify a risk that is amenable to pharmacological intervention. This leads to the conclusion that the NICE assumptions bias cost-effectiveness and unfairly discriminate against women with CRFs.

Side effects

It is arguable whether adverse side effects should be included in the economic assessment of osteoporosis since randomised studies of efficacy have shown few differences between placebo and actively treated patients. NICE elected to incorporate adverse effects, based on a commissioned review of non-randomised studies of the bisphosphonates [100]. It was concluded that women who experience bisphosphonate-related side effects had 91% of the utility of women who do not have such side effects. In the base case analysis for all the drugs under consideration, this was applied to 2.35% of women in the first treatment month and 0.35% of women thereafter. Thus, NICE assumed that the side effect profile of, risedronate, etidronate, raloxifene, strontium and teriparatide were equal, despite evidence to the contrary [100].

In the case of alendronate, the consequences and cost of side effects were assumed by NICE to be ten times greater than that suggested by the systematic review. It was assumed that there would be 23.5 additional GP consultations per 100 patient months in the initial treatment period and 3.5 GP consultations subsequently, and the use of a proton pump inhibitor. Symptoms were assumed to persist for 1 month with a utility loss equivalent to a multiplier of 0.91. The effect of such penalties on cost-effectiveness is marked. With a lifetime horizon, the inclusion of side effects as judged by the review commissioned for NICE had a moderate effect on cost-effectiveness. The ICER increased by 19–40% depending on the clinical scenario (Table 17). As, expected, the effects of using a 10-year rather than a lifetime horizon had a markedly adverse effect on the ICER.

Given the inclusion of adverse side effects, it might be expected that NICE would include side effects that were beneficial to heath. In the case of strontium ranelate, a disease-specific instrument (QUALIOST) showed improvements in quality of life in patients treated with strontium ranelate and a trend in the same direction for a generic instrument (SF-36) [101]. By contrast, NICE assumed that the prevalence and disutility of side effects for strontium ranelate was the same as that assumed by School of Health and Related Research (ScHARR) for the bisphosphonates [100].

In the case of raloxifene, the effect of including non-skeletal effects is even more marked. NICE note that the cost-effectiveness was modelled excluding the risk of venous thromboembolic events (a rare event) and the effect on cardiovascular events (no effect) [102, 103] but still assumed that the prevalence and disutility of side effects for raloxifene was the same as that assumed for the bisphosphonates. However, raloxifene is associated with a marked decrease in breast cancer risk in patients with osteoporosis, with the RR at 4 years for all types of breast cancer reported as 0.38 (95% CI 0.24 to 0.58), and that for invasive breast cancer as 0.28 (95% CI 0.17 to 0.46). The impact of this omission is marked [18] (Table 18). The effect would be greater still in women above the average risk for breast

Table 17 Sensitivity analysis of the cost-effectiveness of alendronate (£/QALY gained) in women aged 70 years [15]

	Time horizon	*T* score=−2.5 SD		No BMD
		No prior fracture	Prior fracture	Prior fracture
Base case (no side effects)	Lifetime	3,709	871	2,130
Base case (no side effects)	10 years	10,950	4,473	7,421
Systematic review	Lifetime	3,780	904	2,172
Systematic review	10 years	11,258	4,604	7,620
Frequency multiplied by 10	Lifetime	4,488	1,222	2,584
Frequency multiplied by 10	10 years	14,796	6,001	9,789

cancer. Whereas NICE conceded the principle that all side effects of using a drug should be considered, it noted that raloxifene was less effective than the bisphosphonates so that adverse effects alone should be included. A further and perhaps more cogent reason given to exclude the effect was that this would require consideration of how raloxifene compared with other drugs that could be used for breast cancer prevention. The comparison is well established [104], but evidently below the radar screen of NICE.

Cost of fracture

Costs of fracture used by NICE were based on long out-dated Health Resource Group codes and are unrealistically low as judged by empirical data in the case of hip fracture, unavailable for vertebral fractures and inappropriate for forearm fractures in the elderly, since a substantial proportion of forearm fractures occur in young individuals [105]. In addition, the incorrect HRG coding was chosen for hip fracture (HRG H85 w/o cc rather than HRG H84 w/cc), decreasing the cost by one third.

These costs contrast with those estimated for the GDG [105] and used to determine cost-effectiveness of intervention in glucocorticoid-induced osteoporosis [59]. Hospital Episode Statistics were used to determine the average length of hospital stay during the period 2002–2004. The estimated cost (Table 19) assumed that the inpatient occupied an orthopaedic bed. These costs, substantially higher than those used by NICE, may even be underestimated and more recent empirical data from the UK suggest even higher costs [106–109]. The more recent study estimated costs at £15,133, £2,753, £1,863, £1,331 and £3,498 for hip, wrist, arm, vertebral and other fractures, respectively.

Cost of drug

The NICE appraisal was undertaken with an assumed annual cost of £53.56 for once-weekly (70 mg) tablets and £108.20 for daily (10 mg) tablets of alendronate and £85.65for etidronate. Both costs were out of date at the time of the FADs and the difference was particularly marked in the case of alendronate. The cost of alendronate used for the appraisal was that given in February 2008, 2 years to the month before the recently revised FADs. Alendronate has subsequently been available at a yearly cost of £26. The NICE appraisals provided no comment on this large decrease in cost, even though the changes invalidate the applicability of the whole appraisal.

Disutility of vertebral fracture

The impact on quality of life the first year after a fracture (hip, vertebral and forearm) was based on empirical estimates [110]. In the case of vertebral fracture, the utility multiplier in the first year was arbitrarily reduced by the appraisal committee by 27% from 0.626 to 0.792, despite empirical evidence to the contrary at the time of the assessment and supported by a systematic review by ScHARR [111]. The issue was not considered to be relevant to the executable model and NICE saw no reason to reverse a decision previously made by the Appraisal Committee. The reason cited was that utility values were based on a hospitalised patient group and not on a typical group of patients with vertebral fractures. As noted by the GDG, the values were, however, not different between hospitalised and out patients [110, 112].

Table 18 Cost-effectiveness of interventions vs. no treatment (£/QALY gained) in women aged 70 years [15]

	T score=−2.5 SD		No BMD
	No prior fracture	Prior fracture	Prior fracture
Alendronate	3,709	871	2,130
Raloxifene	11,184	10,379	10,808
Raloxifene without breast cancer	34,011	23,544	23,755

Table 19 Direct hospital costs of major osteoporotic fractures (£) as used by NICE from Health Resource Group (HRG) codes and those estimated by Stevenson [105] from Hospital Episode Statistics (HES)

Site of fracture	HES	NICE[a]
Hip	10,760	5,157–8,538
Pelvis	9,236	5,157–8,538
Other femoral fractures	13,771	5,157–8,538
Tibia and fibula	3,864	359–585
Spine	1,706	699–803
Humerus	1,112	1,024–1,674
Forearm	527	359–585

[a] Age-dependent range

Discount rates

The costs and the effects in the NICE model were based on discount rates (costs: 6%, effects: 1.5%) that are not in line with the current NICE recommendations (3.5% for both costs and effects). The discount rates are fixed in the model and cannot be changed. NICE acknowledged this, but elected not to remedy this, citing historical precedent, even though the new guideline for NICE appraisals (2004) came into force long before the generation of the 2006 ScHARR model used for the appraisals. Using the older discount rates is likely to underestimate the ICER since the higher discount rates on effects (3.5%) will decrease the QALYs gained.

Willingness to pay

NICE makes the distinction between women who self-identify and those opportunistically assessed. Women who self-identify are 'those that present to a clinician with a clinical risk factor, with no need to find this woman from a multitude of women with the majority having no risk factors. Women could self-identify by having a previous fracture, or reporting one to a clinician, being prescribed glucocorticoids, having a diagnosis of rheumatoid arthritis or consulting a GP concerned about osteoporosis' [55]. In contrast, women who are opportunistically assessed have not presented to a clinician. The distinction is important because the maximum cost per QALY permitted (WTP) was set by NICE at £20,000 per QALY for women who are opportunistically assessed and £30,000 for women who self-identify. NICE considered that women who have already sustained an osteoporotic fracture live with the pain and distress caused by the fracture and that justified a higher WTP. The same consideration was not accorded to women with severe rheumatoid arthritis, ankylosing spondylitis, Crohn's disease, etc.

For the appraisal, a prior fracture alone was used to qualify for self-identification. Thus, those opportunistically assessed would comprise women with a parental history of hip fracture, alcohol intake of 4 or more units per day, and severe long-term rheumatoid arthritis [8, 9]. This seems inequitable given that many women with no prior fracture but with strong risk factors have a fracture risk that exceeds women with a prior fracture but no other risk factors.

Mortality

The NICE appraisals took account of the mortality associated with hip and vertebral fractures by assuming that approximately 42% and 28% of deaths from hip and vertebral fractures, respectively, are causally attributed to the fracture event [113–115]. As noted above, the appraisals, however, did not take account of any mortality consequences associated with the presence or absence of other clinical risk factors.

Conclusion

NICE considers that the effectiveness of interventions in women with CRFs has not been unequivocally demonstrated thus penalising a segment of the osteoporosis community. The same standards are not applied to side effects, the cost of fractures, utilities, mortality and the cost of interventions which has the effect of further penalising cost-effectiveness.

Impact of repopulating the NICE modelling assumptions

For the assumptions reviewed above, NICE concluded that views expressed by consultees on the choice of modelling assumptions including those given by the GDG were not considered to be relevant to the executable model and that there was no reason to reverse decisions previously made by the Appraisal Committee. In many cases, reasons were not given. The effect of accepting the modelling assumptions above may variously over- or underestimate cost-effectiveness, and the question arises as to what is the net effect.

Direct comparisons with the results of NICE are problematic because the NICE evaluations provide estimates of cost-effectiveness over a range of age and range of T score. In addition, detailed ICERs in the context of multiple clinical scenarios were only provided for strontium ranelate. The computation of multiple ICERs was available to consultees for other agents in the form of the ghost model provided by NICE, but the output of this is subject to a confidentiality clause imposed by NICE on the consultees.

We have assessed the impact of the use of FRAX and changing the assumptions on the cost-effectiveness of strontium ranelate using a model that incorporates FRAX

[19]. An intervention for 5 years was modelled. After stopping treatment, the risk reduction was assumed to reverse in a linear manner over a 5-year period as has been used by NICE.

Side effects were not included in the base case since randomised studies of efficacy have shown few persistent differences between placebo and actively treated patients. Indeed, a disease-specific instrument (QUALIOST) showed improvements in quality of life in patients treated with strontium ranelate and a trend in the same direction for a generic instrument (SF-36) [101]. By contrast, NICE assumed that the prevalence and disutility of side effects for strontium ranelate was the same as that assumed by ScHARR for the bisphosphonates [100].

The long-term persistence with strontium ranelate was set at 50% as used in the NICE appraisal. The remaining 50% were assumed to receive 3 months of drug treatment for no health gain [38], as adopted by NICE. A persistence rate of 70% and 30% was assumed for sensitivity analysis (base case±40%).

Costs

Costs of fracture were taken from Stevenson et al. [105] as prepared for the GDG. As noted above, these differed somewhat from those used by NICE, which were based on Health Resource Group codes. Costs did not include any cost for home help. Costs were age-weighted [116] and included nursing home admissions after hip fracture that increased from 6.7% between the age of 50–59 years to 22.6% at the age of 90 years or more [117, 118]. Nursing home costs were not included for fractures at other sites that might require admission to a nursing home.

The cost of medication with strontium ranelate was assumed to be £333.71 per annum (as given in the British National Formulary 54). The cost for case finding was 3 min of GP time to administer the questionnaire on risk factors (£5.76), a BMD test at the femoral neck with dual energy X-ray absorptiometry (£35), and a 10 min consultation with a general practitioner to start treatment (£19.20). Conservatively, all patients treated were assumed to have a BMD test before treatment and 2-yearly thereafter. The unit cost is greater that used in some of the NICE calculations, where identification and assessment costs of treatment were excluded on the presumption that strontium ranelate would be used as a second-line treatment in women not tolerant to alendronate.

Mortality

The age-specific normal mortality rates for the general population in the UK were based on the years 2000–2002. These were adjusted in the model to take into account the mortality associated with the clinical risk factors and any outcome fracture. The NICE appraisals took account of the mortality associated with hip and vertebral fractures by assuming that approximately 30% of deaths are causally attributed to the fracture event [113–115]. The appraisals, however, did not take account of any mortality consequences associated with the presence or absence of other clinical risk factors.

Quality of life

The impact on quality of life the first year after a fracture (hip, vertebral and forearm) was based on empirical estimates [110]. The quality of life estimates for other fractures were based on expert opinion [119]. These multipliers were used together with the population tariff values for the UK [120]. These values are similar to those used by NICE except for vertebral fracture where the utility multiplier in the first year was arbitrarily reduced by the appraisal committee by 27%, despite empirical evidence to the contrary. The effect of this reduction was modelled in sensitivity analyses.

Clinical vignettes

Specific clinical scenarios were used in order to compare as far as possible our estimates of cost-effectiveness with those produced in the NICE appraisals. The clinical risk factors used were those incorporated into the FRAX algorithms developed by the World Health Organization [24, 35, 36]. The clinical risk factors included low BMI (in part dependent on BMD), a prior fragility fracture, a parental history of hip fracture, long-term use (e.g. for 3 months or more) of oral glucocorticoids, rheumatoid arthritis, current cigarette smoking and high alcohol consumption (3 or more units/daily). Fracture probabilities were derived from FRAX®. In contrast, NICE assumed that the weight of each clinical risk factor was equal and used the median estimates of risk.

NICE make the distinction between women who self-identify by having a previous fracture and women who are opportunistically assessed having not presented to a clinician. Those opportunistically assessed comprise women with a parental history of hip fracture, alcohol intake of 4 or more units per day, and severe long-term rheumatoid arthritis [8, 9]. The maximum cost per QALY permitted (WTP) was set at £20,000 per QALY for women who are opportunistically assessed and £30,000 for women who self-identify. For the purpose of this review, we examined the cost-effectiveness of intervention in women with a prior fracture as an example of 'self-identifying' patients (WTP= £30,000) and women with a parental history of hip fracture as an example of 'opportunistic assessment' (WTP=

£20,000). Both scenarios were examined with and without information on BMD. The cost, however, of BMD testing was retained in the examples without information on BMD.

The weight of the various risk factors differs for hip fracture and other fracture outcomes and in the presence or absence of information on BMD. In the absence of a BMD test, BMI is an important predictor of fracture, but is almost entirely dependent on BMD [61]. Thus, for the purposes of modelling, BMI was set to a fixed value of 26 kg/m^2—close to the average value for postmenopausal women.

The incidence of fracture was adjusted to reflect the risk in the target patient groups. The method of calculating fracture risk in the different patient groups relative to the population fracture risks based on BMD and prior fracture is described previously [28] and a similar approach was used for the other risk factors [15]. In addition, the risk of death was adjusted where appropriate according to the presence or absence of the CRFs. Thus, the starting point was the fracture and death hazard in the population with no clinical risk factors and with no BMD test.

For these and other sensitivity analysis, we examined the changes in cost-effectiveness for women at the age of 70 years, as used in an earlier evaluation of alendronate [15].

Results

Table 20 shows that the ICERs in the present study are lower than those of NICE given. In the addendum to the FADs that gives some point estimates for *T* scores and age [66]. Cost-effectiveness ratios were systematically higher in the NICE appraisal than in the present study. Indeed the NICE appraisals gave values of ICER that were up to threefold higher than those in this report. It should be noted that the calculations of NICE did not include acquisition costs (BMD, etc.), whereas BMD testing was included in the present figures. Thus, the differences are likely to be even greater.

The effect of different clinical risk factors at different *T* scores for BMD is shown in Table 21 for women at the age of 70 years [19]. In women at the threshold of osteopenia (a *T* score of −1 SD), treatment with strontium ranelate was cost-effective in the presence of prior fracture or family history. At the threshold of osteoporosis, treatment with strontium ranelate was cost-effective in the presence of any single CRF using a WTP of £30,000, with the exception of current smoking. Prior fractures and a parental history of hip fracture were the strongest risk factors, the use of glucocorticoids and the presence of rheumatoid arthritis had a lesser impact on cost-effectiveness and current smoking and excessive alcohol intake were the weakest of the clinical risk factors.

It is evident that the present model provides cost-effective scenarios for women (with a prior fracture) with *T* scores less than −1 SD, and in opportunistically assessed women at a *T* score between −2.5 and −3.0 SD with a CRF with the exception of smoking. This is in marked contrast to the intervention thresholds derived from the NICE model that ranged from −3.0 to −4.0 SD (see Tables 5 and 6).

In the presence of more than one clinical risk factor, the ICER depended on the weight of the clinical risk factor. In

Table 20 Comparison of estimates of cost-effectiveness of strontium ranelate (£000/QALY gained) in women at specific ages and *T* scores according to the presence or absence of clinical risk factors (CRF; data from [66] and present model as given in [19])

T score	CRFs	Age (years)					
		65		70		75	
		NICE	This report	NICE	This report	NICE	This report
Primary prevention							
−2.5	0	113	36.2	66.3	28.7	52.3	39.1
	1*	94.0	19.4–30.1	55.4	17.5–23.2	43.3	18.4–30.6
−3.0	0	79.2	30.2	48.6	23.8	37.9	19.2
	1*	65.3	17.1–25.1	40.1	14.1–19.0	30.9	11.9–23.8
−3.5	0					26.3	24.7
	1					20.9	9.8–17.9
Secondary prevention							
−3.0	0	47.0	22.4	28.0	17.1	22.4	23.5
	1*	38.4	12.9–18.7	22.8	10.3–13.7	17.9	9.8–17.5
−3.5	0	30.7	18.8	19.1	13.8	15.0	18.2
	1*	23.9	11.4–14.7	15.0	8.9–10.8	11.4	9.4–12.8
Mean value		61.4	21.4	36.9	16.7	23.4	19.1
%NICE			35		45		82

Table 21 Cost-effectiveness of intervention with strontium ranelate (£000/QALY gained) in women aged 70 years with clinical risk factors according to *T* score for femoral neck BMD [19]

	T score (SD)							
	0.0	−0.5	−1.0	−1.5	−2.0	−2.5	−3.0	−3.5
Base case								
Prior fracture	34.2	31.7	29.8	27.6	24.5	20.6	17.1	13.8
Family history	30.9	28.7	26.6	24.3	21.2	17.5	14.2	10.9
Glucocorticoids	36.4	34.6	32.3	30.4	26.6	22.4	18.3	14.0
Rheumatoid arthritis	37.3	34.9	32.5	30.0	26.3	22.0	18.1	14.4
Alcohol >3 units daily	41.1	37.9	35.0	32.2	28.2	23.3	18.9	14.8
Current smoking	62.7	58.4	54.5	50.5	44.3	36.6	29.1	21.8

the absence of information on BMD, the combination of the weakest two risk factors gave an ICER of less than £30,000 (£27,300) at the age of 70 years. In the presence of the strongest two clinical risk factors (family history and prior fracture) and in the absence of information on BMD test, the ICER lay below £20,000/QALY at the age of 70 years (Table 22). In women aged 70 years with a BMD test and two weak CRFs, the ICER was below £30,000/QALY gained with a *T* score of −2.5 SD or less and below £20,000/QALY gained with a *T* score of −3.5 SD or less. With two strong CRFs treatment was cost-effective irrespective of BMD (see Table 6).

Sensitivity analysis

Sensitivity analysis showed that changes in time horizon and assumptions concerning side effects had marked effects on cost-effectiveness (see also Table 17). The ICERs were more than doubled when a 10-year rather than a lifetime horizon was used. When side effects, as assumed by the systematic review for NICE, were included, this had a lesser, though marked effect on cost-effectiveness using the lifetime horizon, but had a more marked adverse effect on cost-effectiveness with the shorter time horizon (see Table 17). Moderate effects on cost-effectiveness were observed with changes in the assumptions concerning offset time, adherence, and utility weights for spine fracture (Table 23).

Table 22 Cost-effectiveness of treatment with strontium ranelate in women aged 70 years and two weak clinical risk factors (current smoking and excessive alcohol intake) or two strong risk factors (family history and prior fracture)

BMD *T* score	Base case	
	Weak CRFs	Strong CRFs
No BMD	−	++
−1.0	−	++
−1.5	−	++
−2.0	−	++
−2.5	+	++
−3.0	+	++
−3.5	++	++

++ denotes cost-effectiveness at a WTP of £30,000, + denotes cost-effectiveness at a WTP of £20,000, − cost-ineffective

Intervention thresholds

A strength of FRAX® is the ability to express risk as fracture probabilities which are more readily understood than *T* scores by physicians and patients. The inappropriateness of the use of a single *T* score to direct intervention is now widely acknowledged. Thus probability-based assessment is now becoming the norm for treatment guidelines [4, 23, 24, 33, 36, 41, 42, 70–72, 75, 76, 79, 81, 88, 121–123]. The FRAX® models were supplied to avail NICE of the opportunity to use probabilities as intervention thresholds as recommended by the Guideline Development Group of NICE.

NICE have been reluctant to adopt this approach for three reasons, which are variously ill-founded or spurious [8, 9, 12, 13]. It is said that NICE did not have access to the FRAX® algorithms. This, however, is misleading since they were offered FRAX (in confidence) 5 years ago to the Assessment Group. Unfortunately, the data were incompletely used and in some instances inappropriately used.

A second reason given by NICE for avoiding probability-based treatment thresholds in favour of those based on *T* scores for BMD is the argument that absolute fracture risk does not provide a single measure of cost-effectiveness. This is correct as shown in Table 16 and elsewhere [55]. For example, at a WTP of £30,000, it was cost-effective on average to intervene with strontium ranelate in a woman aged 70 years with a *T* score of −2.5 SD and a clinical risk factor [19]. From the different permutations of risk factors the ICER might range from £17,500 to £23,200/QALY gained, i.e. a 1.3-fold range in cost-effectiveness ratio. The argument is spurious given that the use by NICE of *T* score ranges, median weights for CRFs and age ranges for their economic assessment give an equivalent or, more usually, a greater variance.

Table 23 Sensitivity analysis of the cost-effectiveness of strontium ranelate in women aged 70 years [19]

	Cost (£000)/QALY gained				
		Self-identifying		Opportunistic case finding	
	T score=−2.5	*T* score=−2.5	No BMD	*T* score=−2.5	No BMD
	No previous fracture	+ previous fracture	+ previous fracture	+ parental history	+ parental history
Base case	28.7	20.6	21.5	17.5	22.5
NICE efficacy	29.2	20.9	21.8	17.9	22.9
10 year time horizon	64.3	46.9	53.1	36.8	52.8
Offset time +40% (7 years)	26.5	18.9	19.7	15.9	20.6
Offset time −40% (3 years)	31.5	22.6	23.6	19.4	24.8
Non-adherence +40% (70%)	30.8	22.9	23.9	18.9	24.1
Non-adherence −40% (30%)	27.8	19.5	20.4	16.8	21.7
Higher utility for vertebral fracture	30.4	22.1	22.8	18.5	23.6
Side effects	46.3	28.2	29.6	23.2	32.1

A third reason given by NICE is that the Appraisal Committee 'was not persuaded that the drugs under consideration had been unequivocally shown to reduce fracture risk that was attributable to risk factors not mediated through low BMD and age.' This is ironic given that NICE vary intervention thresholds according to the number of risk factors and even accord less efficacy to intervention in the presence of risk factors. More disturbing is that the view is not in accordance with the available evidence as reviewed in this paper and elsewhere [15, 24, 42, 52, 53, 68].

A final argument, raised in 2010, was that the consideration of intervention thresholds on the basis of fracture probability would require a new appraisal, a view expressed by the GDG and other consultees in 2005.

Indeed, it is relatively straightforward to produce intervention thresholds based on the probability of fracture [42]. Intervention thresholds at each age can be determined from the relationship between fracture probabilities (clinical spine, hip, forearm or humerus) and the cost-effectiveness of all possible combinations of CRFs at BMD *T* scores between 0 and −3.5 SD in 0.5 SD steps (512 combinations) with a BMI set to 26 kg/m^2. Note that this approach is not a population simulation, but an array of all possible combinations.

At each age, there was a close correlation between the probability of a major osteoporotic fracture as determined by FRAX® and cost-effectiveness. The relationship is illustrated in Fig. 4 for women at the age of 50 years treated with alendronate [42].

The point estimates for the correlations permit the calculation of the mean fracture probability for any willingness to pay as shown in Table 24 for a WTP of £20,000. There was rather little difference in the threshold probability at which treatment became cost-effective at different ages with a mean value of 6.9% at a WTP of £20,000. Thus, with a WTP of £20,000, any recommendations for intervention should ensure that individuals have a fracture probability that exceeds 7%. The NOGG guidance outlined previously (see Fig. 1) succeeds in this expectation.

Probability-based thresholds will differ according to the effectiveness and cost of intervention. For risedronate [20], as noted for alendronate, there was little difference in the threshold probability at which treatment became cost-effective at different ages with a mean value of 18.6% at a WTP of £20,000 and 13.0% and at a WTP of £30,000 (Table 25). Intervention thresholds have also been determined for strontium ranelate [19]. The mean threshold values were a 10-year probability of 21.6% with a WTP of £30,000 and 37.8% with a WTP of £20,000.

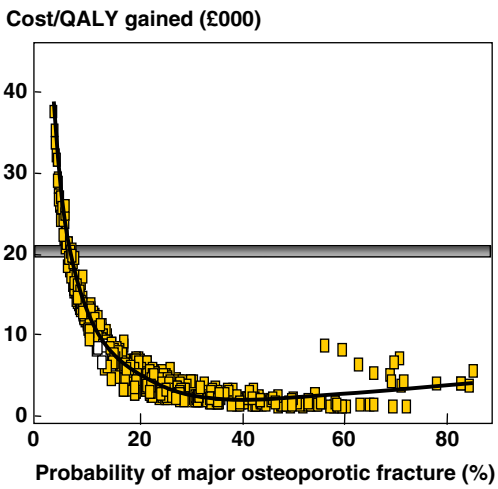

Fig. 4 Correlation between the 10 year probability of a major osteoporotic fracture and cost-effectiveness of alendronate at the age of 50 years in women (BMI set to 26 kg/m^2; each point represents a particular combination of clinical risk factors) [42]

Table 24 Ten-year probabilities (mean and 95% confidence intervals; CI) of a major osteoporotic fracture (%) by age at or above which treatment with alendronate becomes cost-effective [42]

Age (years)	10 year probability of osteoporotic fracture (%)	
	£20,000/QALY	
	Probability	95% CI
50	5.6	4.8–6.8
55	7.6	6.3–8.9
60	8.4	7.1–9.6
65	6.1	4.9–7.2
70	5.0	4.2–5.9
75	7.3	6.6–8.1
80	8.3	7.8–8.8
Mean	6.9	

Discussion

The translation of the NICE appraisals for the assessment and treatment of osteoporosis is difficult—to say the least. The intervention thresholds are complex and ill-suited to primary care. Can we really live with different treatment thresholds for different interventions where women who start treatment on alendronate but are unable to tolerate it will have to wait for their disease to progress before they can receive another treatment? How can general practitioners explain to women in whom alendronate is contraindicated that they cannot be given alterative treatment despite being at high risk of fracture and the availability of effective and relatively cheap alternatives such as raloxifene, strontium ranelate and risedronate? How will the general practitioner justify to carers and patients the discrimination against the disabled and the frail elderly populations in whom alendronate is contraindicated because of cognitive dysfunction (and therefore inability to comply with the dosing instructions) or physical frailty?

Notwithstanding these ethical considerations, the complexity of the recommendations for alternative interventions makes them clinically unworkable. The physician is left with no advice for women at high risk who do not also have a BMD *T* score of less than −2.5 SD and a prior fragility fracture. The problem is compounded by the fact that NICE ignored the consistent advice of the GDG and the National Osteoporosis Society that recommendations about treatment be based on 10-year fracture probability, as in the FRAX algorithm, rather than on *T* scores, age and the number of risk factors. FRAX is widely available (http://www.shef.ac.uk/ FRAX) and increasingly used in clinical practice worldwide, including the UK. No guidance whatsoever is provided for men, for men or women on treatment with glucocorticoids or for many of the other secondary causes of osteoporosis and newer treatments such as ibandronate and zoledronic acid are not included.

These problems have arisen because of dysfunctional consultation process where advice has gone largely unheeded. Indeed, the GDG, whose responsibility is to turn guidance into guidelines, has long been suspended and representation of the GDG at Committee meetings denied even before the suspension. The end result is guidance that serves neither primary care physicians nor the patients at high risk of fracture.

This problem is compounded by concerns surrounding the technology appraisals for osteoporosis. A major difficulty has been the lack of transparency in the model construct and the manner in which the model has been populated. A stumbling block was the reluctance on the part of NICE to negotiate the restricted release of confidential information supplied to NICE by the WHO Collaborating Centre for Bone Diseases at Sheffield. This impediment is now resolved, but the problems remain.

The model supplied for the consultation period is opaque. It is based on an individual state transition model. The authors claim that individual patient simulations are superior to cohort models in the accuracy with which they are populated and their flexibility. It is ironic that the model construct appears to be inflexible in that, rather than rebuild

Table 25 Probability (mean and 95% confidence intervals; CI) within 10 years of a major osteoporotic fracture (%) by age at or above which treatment with risedronate becomes cost-effective [20]

Age (years)	10-year probability of osteoporotic fracture (%) with BMD at a WTP of			
	£20,000/QALY		£30,000/QALY	
	Probability	95% CI	Probability	95% CI
50	17.1	12.1–29.4	10.0	8.6–14.9
55	19.8	16.1–29.3	14.3	10.6–18.0
60	23.0	17.5–33.2	16.5	12.6–20.8
65	18.0	14.4–23.9	11.9	9.4–15.4
70	16.1	12.9–19.2	9.9	8.9–12.9
75	17.9	13.8–23.3	13.3	9.7–17.4
80	18.3	14.7–24.3	15.2	11.6–18.9

the model to fit the requirements of the NICE appraisal, the model has been successively adapted with transformations and add-ons which makes it susceptible to accuracy errors. The adaptations are so extensive that it is quite uncertain whether it can still be defined as a state transition model any longer. Indeed the model supplied is a 'model of a model' or as we have termed it, a ghost model.

Unfortunately, the manner whereby these transformations and add-ons were computed and the assumptions used are for the most part opaque and supplied in neither the HTA reports nor the Appraisal documentation. This was only partly redressed in the DSU report but the details remain confidential. Thus, we consider that a fully transparent model was not supplied for evaluation. This is a matter for concern. The concern is accentuated by the observation that, where we were able to deconstruct the model, in very few instances could we replicate the findings of the authors of the model.

The lack of transparency has a number of consequences. Firstly, the model cannot be externally validated. Secondly, internal validation is problematic. Indeed a component of this review has been to address its internal validation. Thirdly, the transformations have meant that many variables necessary for sensitivity testing cannot be accessed or varied. As a trivial example, it is not even possible to model changes in the discount rates for QALYs and costs, with the result that the rates used do not conform to those recommended by NICE.

A potential strength of the model is the incorporation of the FRAX algorithms. This permits the use of multiple risk factors for fracture (and death) to be integrated for the assessment of fracture probability. The obvious application of FRAX is in the assessment of individuals to identify those who would be candidates for pharmacological intervention, and it has been widely used since the launch of the web site, currently receiving on average 200,000 hits daily. Unfortunately, NICE did not use this methodology for the computation of 10-year probability or for the computation of intervention thresholds. Indeed, it is not even possible to compute 10-year fracture probability from the model. The arguments of NICE for not basing intervention thresholds on fracture probabilities are not robust, and indeed intervention thresholds for several treatments are given within this review.

These considerations apart, a second reason for NICE to use the information provided by FRAX was to improve the accuracy in stratifying risk and therefore cost-effectiveness. Indeed it is integral to the NICE model. Since the FRAX variables were supplied by one of the authors, the manner in which the information has been used has in some measure been more readily deconstructed than most of the other add-ons or transformations. The use made of FRAX is problematic. The death hazards are ignored or inappropriately used, continuous variables ignored, the categorisation of risk factors changed, and risk factors inappropriately used for costing an identification strategy. In addition, numerous errors have been identified, each of which may be minor, but compounded, have an uncertain effect on accuracy. It is unfortunate that accuracy could not be directly tested since outputs of the modelled probability are not supplied.

Several health economic assessments have drawn attention to discrepancies in estimates of cost-effectiveness produced by NICE and other models [12, 13, 16, 18–20, 124, 125]. It is difficult to determine why the results differ, but ultimately reasons reside in the construct of the model, the assumptions used to populate the model or both. With regard to construct, the NICE model uses predominantly a 10-year time horizon, which, as shown in previous sensitivity analyses and in this review has a large effect on apparent cost-effectiveness. The adequacy of the add-ons was impossible to address due to the lack of transparency of the model.

In order to address this issue we built a 'replica' model. One conclusion from the replica model was that a large component of the difference in cost-effectiveness resided in the assumptions used to populate the model. The assumptions that we used were based on empirical observation rather than expert opinion as detailed in this review. We cannot report on the difference in results between the authentic NICE model and the replica model for reasons of confidentiality but there were systematic and non-systematic differences when the replica model was populated with the same assumptions as the NICE model. The NICE appraisal confirms discrepancies between the model of NICE and the replica model (referred to as the results being 'largely similar'). The term 'largely similar' refers to the much greater differences when the replica model was populated with empirical data rather than by expert opinion. In addition, the numerous errors found in the accessible parts of the model impair significantly the stratification of risk and thus the effective targeting of treatment.

It is unfortunate that the NICE could not or would not consider comments on the effect of different values of input parameters on model outputs (as these had been apparent in the initial appraisal) or comments made on aspects of the model that had previously been described in Assessment Reports or other consultation documents. This has led to rejecting or ignoring empirical data, preferring to defer to expert opinion. This does not lend itself to a claim that cost-effective analysis is evidence-based. Indeed, the reason why recommendations for alendronate have not changed in 5 years despite the sixfold reduction in price include a reduction in the efficacy estimate for alendronate, reduction in the disutility value for vertebral fractures, reduction in the cost per QALY threshold for primary prevention and a

new assumption of the frequency of side effects that was then multiplied by an arbitrary factor of 10. All these changes have been made in the absence of any change in the evidence base.

This apart, the conclusion of NICE that there are no issues that have been raised by consultees which cause it to doubt the validity of the model or that raise justifiable doubts about the appropriateness of the use of the model to inform the guidance seem at best to be very overoptimistic.

The osteoporosis community recognises the need to use healthcare resources effectively—and only asks that osteoporosis receives an appraisal that is equitable compared with other chronic diseases. The evidence suggests that osteoporosis can compete both in terms of the burden of disease and health economics [126, 127]. Just as a prior fracture is a strong risk factor for a further fracture, is the failure of NICE to serve osteoporosis over very many years a sign that it will continue to do so? It is perhaps no surprise that the vacuum created by NICE and the availability of the FRAX tools has stimulated the development of clinical guidelines in many countries, including the UK, so that the needs of patients can be met. Perhaps NICE should be thanked for this.

Acknowledgements We are grateful to Professors Juliet Compston and Cyrus Cooper for their review of the manuscript and constructive comments.

Conflicts of interest This review is an adaptation of a more extensive report commissioned by Servier for use in their response to consultation documents issued by NICE in 2009. Servier Laboratories provided financial support for the original report. Servier had no role in study design, in the collection, analysis, and interpretation of data and in the writing of the report or this review. JAK, BJ, EM, AC, OS and FB act as advisors to and have received funding from many pharmaceutical companies involved in marketing products for treatment of osteoporosis. JAK was a member of the NICE Osteoporosis GDG, and was asked by NICE to leave the GDG in 2006. He is also a member of NOGG. JAK is President of the International Osteoporosis Foundation and FB serves on its Committee of Scientific Advisors. AC is a member of NOGG. OS collaborates with the Osteoporosis Guideline Group of the Swedish National Board of Health and Welfare.

References

1. Kanis JA, Johnell O (2005) Requirements for DXA for the management of osteoporosis in Europe. Osteoporos Int 16:229–238
2. Dolan P, Torgerson DJ (2000) The cost of treating osteoporotic fractures in the United Kingdom female population. Osteoporos Int 11:551–552
3. Delmas PD (2002) Treatment of postmenopausal osteoporosis. Lancet 359:2018–2026
4. Kanis JA, Burlet N, Cooper C, Delmas PD, Reginster J-Y, Borgstrom F, Rizzoli R, on behalf of the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO) (2008) European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int 19:399–428
5. National Institute for Clinical Excellence (2005) Bisphosphonates (alendronate, etidronate, risedronate), selective oestrogen receptor modulators (raloxifene) and parathyroid hormone (teriparatide) for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
6. National Institute for Health and Clinical Excellence (2007) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
7. National Institute for Health and Clinical Excellence (2007) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
8. National Institute for Health and Clinical Excellence (2008) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
9. National Institute for Health and Clinical Excellence (2008) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
10. The Queen (2009) Case number C1/2009/0805 The Queen (on the application of Servier Laboratories Ltd) v The National Institute for Health and Clinical Excellence
11. The Queen (2009) (on the application of Servier Laboratories Ltd) v The National Institute for Health and Clinical Excellence. EWHC 281 (Admin)
12. National Institute for Health and Clinical Excellence (2010) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
13. National Institute for Health and Clinical Excellence (2010) Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
14. Delmas PD, Siris ES (2008) NICE recommendations for the prevention of osteoporotic fractures in postmenopausal women. Bone 42:16–18
15. Kanis JA, Adams J, Borgström F et al (2008) Cost-effectiveness of alendronate. Bone 42:4–15
16. Kanis JA, Adams J, Borgström et al (2008) Modelling cost-effectiveness for osteoporosis. Bone 43:215–216
17. Kanis JA, Compston JE, National Osteoporosis Guideline Group of the UK (2008) NICE continues to muddy the waters of osteoporosis. Osteoporos Int 19:1105–1107
18. Kanis JA, Borgstrom F, Johnell O, Oden A, Sykes D, Jonsson B (2005) Cost-effectiveness of raloxifene in the UK. An economic evaluation based on the MORE study. Osteoporos Int 16:15–25
19. Borgström F, Ström O, Coelho J, Johansson H, Odén A, McCloskey E, Kanis JA (2010) The cost-effectiveness of strontium ranelate in the UK for the management of osteoporosis. Osteoporos Int 21:339–349
20. Borgström F, Ström O, Coelho J, Johansson H, Oden A, McCloskey EV, Kanis JA (2010) The cost-effectiveness of risedronate in the UK for the management of osteoporosis using the FRAX. Osteoporos Int 21:495–505
21. Mills J (2009) Servier wins review of NICE guidance on Protelos. Scrip 3434 Feb 27th p1 and 3
22. National Collaborating Centre for Nursing and Supportive Care (2008) Systematic reviews of clinical effectiveness prepared for

the guideline. Osteoporosis: assessment of fracture risk and the prevention of osteoporotic fractures in individuals at high risk National Institute for Health and Clinical Excellence. (http://www.nice.org.uk/nicemedia/pdf/OsteoporosisEvidenceReviews190908.pdf). Posted September 2008, Accessed 25th March 2010
23. National Osteoporosis Guideline Group on behalf of the Bone Research Society, British Geriatrics Society, British Society of Rheumatology, Society of Endocrinology, British Orthopaedic Association, Primary Care Rheumatology Society, Osteoporosis 2000 and Osteoporosis Dorset (2008) Osteoporosis: clinical guideline for prevention and treatment. University of Sheffield Press, Sheffield
24. Kanis JA, on behalf of the World Health Organization Scientific Group (2008) Assessment of osteoporosis at the primary health-care level. Technical report. University of Sheffield, UK: WHO Collaborating Center; 2008.Available at http://www.shef.ac.uk/FRAX/index.htm
25. Royal College of Physicians (1999) Osteoporosis: clinical guidelines for the prevention and treatment. Royal College of Physicians, London
26. Royal College of Physicians and Bone and Tooth Society of Great Britain (2000) Update on pharmacological interventions and an algorithm for management. Royal College of Physicians, London
27. Kanis JA, Delmas P, Burckhardt P, Cooper C, Torgerson D (1997) Guidelines for diagnosis and management of osteoporosis. The European Foundation for Osteoporosis and Bone Disease. Osteoporos Int 7:390–406
28. Kanis JA, Torgerson D, Cooper C (2000) Comparison of the European and USA practice guidelines for Osteoporosis. Trends Endocrinol Metab 11:28–32
29. Kanis JA (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. WHO Study Group. Osteoporos Int 4:368–381
30. Schuit SC, van der Klift M, Weel AE et al (2004) Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. Bone 34:195–202
31. Wainwright SA, Marshall LM, Ensrud KE, Cauley JA, Black DM, Hillier TA, Hochberg MC, Vogt MT, Orwoll ES, Study of Osteoporotic Fractures Research Group (2005) Hip fracture in women without osteoporosis. J Clin Endocrinol Metab 90:2787–2793
32. Black DM, Steinbuch M, Palermo L et al (2001) An assessment tool for predicting fracture risk in postmenopausal women. Osteoporos Int 12:519–528
33. Kanis JA (2002) Diagnosis of osteoporosis and assessment of fracture risk. Lancet 359:1929–1936
34. Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J et al (2007) The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. Osteoporos Int 18:1033–1046
35. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey EV (2008) FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19:385–397
36. World Health Organization (2007) Assessment of osteoporosis at the primary health care level. WHO, Geneva, (www.who.int/chp/topics/rheumatic/en/index.html)
37. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP (1998) Updated data on proximal femur bone mineral levels of US adults. Osteoporos Int 8:468–486
38. Kanis JA, Johnell O, Oden A et al (2002) Ten-year risk of osteoporotic fracture and the effect of risk factors on screening strategies. Bone 30:251–258
39. De Laet C, Oden A, Johansson H et al (2005) The impact of the use of multiple risk indicators for fracture on case-finding strategies: a mathematical approach. Osteoporos Int 16:313–318
40. Johansson H, Kanis JA, Oden A, Johnell O, McCloskey E (2009) BMD, clinical risk factors and their combination for hip fracture prevention. Osteoporos Int 20:1675–1682
41. Compston J, Cooper A, Cooper C, Francis R, Kanis JA, Marsh D, McCloskey EV, Reid DM, Selby P, Wilkins M, on behalf of the National Osteoporosis Guideline Group (NOGG) (2009) Guidelines for the diagnosis and management of osteoporosis in postmenopausal women and men from the age of 50 years in the UK. Maturitas 62:105–108
42. Kanis JA, McCloskey EV, Johansson H, Strom O, Borgstrom F, Oden A, the National Osteoporosis Guideline Group (2008) Case finding for the management of osteoporosis with FRAX® - Assessment and intervention thresholds for the UK. Osteoporos Int 19: 1395-1408 Erratum 2009. Osteoporos Int 20:499–502
43. Johansson H, Kanis JA, Oden A, Johnell O, Compston J, McCloskey E (2010) A comparison of case finding strategies for the management of osteoporosis. Osteoporos Int. accepted for publication, September 2010
44. Kanis JA, Borgstrom F, Johnell O, Jonsson B (2004) Cost-effectiveness of risedronate for the treatment of osteoporosis and prevention of fractures in postmenopausal women. Osteoporos Int 15:862–871
45. Fleurence RL, Iglesias CP, Johnson JM (2007) The cost effectiveness of bisphosphonates for the prevention and treatment of osteoporosis: a structured review of the literature. Pharmacoeconomics 25:913–933
46. Iglesias CP, Torgerson DJ, Bearne A, Bose U (2002) The cost utility of bisphosphonate treatment in established osteoporosis. Q J Med 95:305–311
47. Borgström F, Ström O, Kanis JA (2008) What proportion of patients can be treated with second line at a maintained incremental cost utility ratio. Osteoporos Int 19(Suppl 1):S65
48. Mitchell PJ and Bayly JR (2009) A comparative analysis of drug prescribing and costs for osteoporosis and hypercholesterolaemia in England for 1998–2007. J Bone Miner Res 24 (Suppl 1). Available at http://www.asbmr.org/Meetings/AnnualMeeting/AbstractDetail.aspx?aid=469d76bd-613b-4262-869b-59d11dd16fa6. Accessed March 16, 2010
49. Cummings SR, Black DM, Thompson DE et al (1998) Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA 280:2077–2082
50. Black DM, Cummings SR, Karpf DB et al (1996) Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet 348:1535–1541
51. McClung MR, Geusens P, Miller PD et al (2001) Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Engl J Med 344:333–340
52. McCloskey EV, Johansson H, Oden A et al (2009) Ten-year fracture probability identifies women who will benefit from clodronate therapy—additional results from a double blind, placebo controlled randomised study. Osteoporos Int 20:811–818
53. Kanis JA, Johansson H, Oden A, McCloskey EV (2009) Bazedoxifene reduces vertebral and clinical fractures in post-menopausal women at high risk assessed with FRAX®. Bone 44:49–54
54. Stevenson M, Lloyd Jones M, De Nigris E, Brewer N, Davis S, Oakley J (2005) A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis. Health Technol Assess 9:1–160
55. Stevenson M, Davis S, Lloyd-Jones M, Beverley C (2007) The clinical effectiveness and cost-effectiveness of strontium ranelate for the prevention of osteoporotic fragility fractures in postmenopausal women. Health Technol Assess 11:1–134

56. Stevenson MD, Oakley J, Chilcott JB (2004) Gaussian process modeling in conjunction with individual patient simulation modeling: a case study describing the calculation of cost-effectiveness ratios for the treatment of established osteoporosis. Med Decis Mak 24:89–100
57. Stevenson MD, Brazier JE, Calvert NW, Lloyd-Jones M, Oakley J, Kanis JA (2005) Description of an individual patient methodology for calculating the cost-effectiveness of treatments for osteoporosis. J Oper Res Soc 56:214–221
58. Zethraeus N, Borgstrom F, Strom O, Kanis JA, Jonsson B (2007) Cost-effectiveness of the treatment and prevention of osteoporosis—a review of the literature and a reference model. Osteoporos Int 18:9–23
59. Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M (2007) Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis. Health Technol Assess 11:1–256
60. Ström O, Borgström F, Kanis JA, Jönsson B (2009) Incorporating adherence in health economic modelling of osteoporosis. Osteoporos Int 20:23–34, with erratum page 35
61. De Laet C, Kanis JA, Oden A, Johanson H, Johnell O, Delmas P et al (2005) Body mass index as a predictor of fracture risk: a meta-analysis. Osteoporos Int 16:1330–1338
62. Kanis JA, Johansson H, Johnell O et al (2005) Alcohol intake as a risk factor for fracture. Osteoporos Int 16:737–742
63. Dam TT, Harrison S, Fink HA, Ramsdell J, Barrett-Connor E, for the Osteoporotic Fractures in Men (MrOS) Research Group (2010) Bone mineral density and fractures in older men with chronic obstructive pulmonary disease or asthma. Osteoporos Int 21:1341–1349
64. Kanis JA, Johnell O, Oden A et al (2005) Smoking and fracture risk: a meta-analysis. Osteoporos Int 16:155–162
65. Wong PK, Christie JJ, Wark JD (2007) The effects of smoking on bone health. Clin Sci (Lond) 113:233–241
66. National Institute for Health and Clinical Excellence (2008) Addendum to Final appraisal determination. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the primary/secondary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, London
67. Praities N (2009) Government backs rival to NICE guidance on fracture risk. Pulse. Saturday, 15 August 2009
68. Kanis JA, Oden A, Johansson H, Borgström F, Ström O, McCloskey E (2009) FRAX® and its applications to clinical practice. Bone 44:734–743
69. National Osteoporosis Foundation (2008) Clinician's guide to prevention and treatment of osteoporosis. Washington, DC: National Osteoporosis Foundation. www.nof.org
70. Neuprez A, Johansson H, Kanis JA et al (2009) Rationalisation du remboursement des médicaments de l'ostéoporose: de la mesure isolée de la densité osseuse à l'intégration des facteurs cliniques de risque fracturaire. Validation de l'algorithme FRAX®. Rev Méd Liège 64(12):612–619
71. Fujiwara S, Nakamura T, Orimo H, Hosoi T, Gorai I, Oden A et al (2008) Development and application of a Japanese model of the WHO fracture risk assessment tool (FRAX™). Osteoporos Int 19:429–448
72. Lippuner K, Johansson H, Kanis JA, Rizzoli R (2010) FRAX® assessment of osteoporotic fracture probability in Switzerland. Osteoporos Int 21:381–390
73. Czerwinski E, Kanis JA, Trybulec B, Johansson H, Borowy P, Osieleniec J (2009) The incidence and risk of hip fracture in Poland. Osteoporos Int 20:1363–1368
74. Berry SD, Kiel DP, Donaldson MG, Cummings SR, Kanis JA, Johansson H, Samelson EJ (2010) Application of the National Osteoporosis Foundation Guidelines to postmenopausal women and men: the Framingham Osteoporosis Study. Osteoporos Int 21:53–60
75. Siminoski K, Leslie WD, Frame H, Hodsman A, Josse RG, Khan A, Lentle BC et al (2007) Recommendations for bone mineral density reporting in Canada: a shift to absolute fracture risk assessment. J Clin Densitom 10:120–123
76. Dawson-Hughes B, Tosteson AN, Melton LJ 3rd, Baim S, Favus MJ, Khosla S et al (2008) Implications of absolute fracture risk assessment for osteoporosis practice guidelines in the USA. Osteoporos Int 19:449–458
77. Dawson-Hughes B, Looker AC, Tosteson ANA, Johansson H, Kanis JA, Melton LJ III (2010) The potential impact of new National Osteoporosis Foundation guidance on treatment patterns. Osteoporos Int 21:41–52
78. Tsang SWY, Kung AWC, Kanis JA, Johansson H, Odén A (2009) Ten-year fracture probability in Hong Kong southern Chinese according to age and BMD femoral neck T-scores. Osteoporos Int 20:1939–1945
79. Kurth AA, Pfeilschifter J (2007) Diagnosis and treatment of postmenopausal osteoporosis and osteoporosis in men. German Guidelines Update 2006. Orthopade 36:683–690, German
80. Leslie WD, Lix LM, Johnansson H, Odén A, McCloskey E, Kanis JA (2010) Independent clinical validation of a Canadian FRAX tool: Fracture prediction and model calibration. J Bone Miner Res, in press PMID: 20499367
81. Adami S, Bertoldo F, Brandi ML et al (2009) Guidelines for the diagnosis, prevention and treatment of osteoporosis. Reumatismo 61:260–284, Italian
82. Strom O, Macarios D, Badamgarav E, Borgstrom F, Tosteson A, Kanis J (2009) A UK denosumab cost-effectiveness model incorporating FRAX and adherence. J Bone Miner Res 24(Suppl 1):s141
83. Borgström F, Ström O, Kleman M, McCloskey EV, Johansson H, Odén A, Kanis JA (2010) Cost-effectiveness of bazedoxifene incorporating the FRAX® algorithm in a European perspective. Osteoporos Int, in press PMID: 20532482
84. Johansson H, Oden A, Johnell O, Jonsson B, De Laet C, Oglesby A et al (2004) Optimisation of BMD measurements to identify high risk groups for treatment—a test analysis. J Bone Miner Res 19:906–913
85. Kanis JA, Johnell O, De Laet C et al (2004) A meta-analysis of previous fracture and subsequent fracture risk. Bone 35:375–382
86. European Community (1998) Report on osteoporosis in the European Community. 1998. EC, Strasbourg
87. National Osteoporosis Foundation (NOF) (2003) Physician's guide to prevention and treatment of osteoporosis. National Osteoporosis Foundation, Washington
88. Committee for Medicinal Products for Human Use (CHMP) (2006) Guideline on the evaluation of medicinal products in the treatment of primary osteoporosis. Ref CPMP/EWP/552/95Rev.2. CHMP, London
89. Saag KG, Emkey R, Schnitzer TJ, Brown JP, Hawkins F, Goemaere S et al (1998) Alendronate for the prevention and treatment of glucocorticoid-induced osteoporosis. Glucocorticoid-induced Osteoporosis Intervention Study Group. New Engl J Med 339:292–299
90. Adachi JD, Saag KG, Delmas PD, Liberman UA, Emkey RD, Seeman E et al (2001) Two year effects of alendronate on bone mineral density and vertebral fracture in patients receiving glucocorticoids: a randomised, double-blind, placebo-controlled extension trial. Arthritis Rheum 44:202–211
91. Reginster JY, Minne HW, Sorensen OH et al (2000) Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral efficacy with risedronate therapy (VERT) study group. Osteoporos Int 11:83–91
92. Roussow JE, Anderson GL, Prentice RL et al (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal

women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 288:321–333
93. McCloskey EV, Beneton M, Charlesworth D et al (2007) Clodronate reduces the incidence of fractures in community dwelling elderly women unselected for osteoporosis: results of a double-blind, placebo-controlled randomized study. J Bone Miner Res 22:135–141
94. Chapuy MC, Arlot ME, Delmas PD, Meunier PJ (1994) Effect of calcium and cholecalciferol treatment for three years on hip fractures in elderly women. BMJ 308:1081–1082
95. Kanis JA, Barton I, Johnell O (2005) Risedronate decreases fracture risk in patients selected solely on the basis of prior vertebral fracture. Osteoporos Int 16:475–482
96. McCloskey EV, Selby P, Davies M et al (2004) Clodronate reduces vertebral fracture risk in women with post-menopausal or secondary osteoporosis: results of a double blind placebo-controlled 3 year study. J Bone Miner Res 19:728–736
97. Roux C, Reginster J-Y, Fechtenbaum J et al (2006) Vertebral fracture with reduction with strontium ranelate in women with postmenopausal osteoporosis is independent of baseline risk factors. J Bone Miner Res 21:536–542
98. Marcus R, Wang O, Satterwhite J et al (2003) The skeletal response to teriparatide is largely independent of age, initial bone mineral density and prevalent vertebral fractures in postmenopausal women with osteoporosis. J Bone Miner Res 18:18–23
99. Johnell O, Kanis JA, Black DM et al (2004) Association between baseline risk factors and vertebral fracture risk in the Multiple Outcomes of Raloxifene Evaluation (MORE) study. J Bone Miner Res 19:764–772
100. Lloyd Jones M, Wilkinson A (2006) Adverse effects and persistence with therapy in patients taking oral alendronate, etidronate or risedronate: a systematic review. NHS R & D HTA. ScHARR
101. Marquis P, Roux C, de la Loge C, Diaz-Curiel M, Cormier C, Isaia G, Badurski J, Wark J, Meunier PJ (2008) Strontium ranelate prevents quality of life impairment in post-menopausal women with established vertebral osteoporosis. Osteoporos Int 19:503–510
102. Collins P, Mosca L, Geiger MJ et al (2009) Effects of the selective estrogen receptor modulator raloxifene on coronary outcomes in the Raloxifene Use for The Heart trial: results of subgroup analyses by age and other factors. Circulation 24:922–930
103. Barrett-Connor E, Cox DA, Song J, Mitlak B, Mosca L, Grady D (2009) Raloxifene and risk for stroke based on the Framingham stroke risk score. Am J Med 122:754–761
104. Nelson HD, Fu R, Griffin JC, Nygren P, Smith ME, Humphrey L (2009) Systematic review: comparative effectiveness of medications to reduce risk for primary breast cancer. Ann Intern Med 151:703–715, W-226-35
105. Stevenson M, Davis SE, Kanis J (2006) The hospitalization costs and outpatient costs of fragility fractures. Women's Health Med 4:149–151
106. Lawrence TM, White CT, Wenn R, Moran CG (2005) The current hospital costs of treating hip fractures. Injury 36:88–91, discussion 92
107. Johal K, Boulton C, Sahota O, Moran C (2007) Cost versus funding for hip fracture treatment: financial suicide for NHS Trusts? Osteoporos Int 18(Suppl 3):S245–S328
108. Iglesias CP, Manca A, Torgerson DJ (2009) The health-related quality of life and cost implications of falls in elderly women. Osteoporos Int 20:869–878
109. World Health Organisation (2007) Falls prevention in older age. WHO Press, Geneva (www.who.int/ageing/publications/Falls_prevention7March.pdf). Accessed 15 January 2010
110. Borgström F, Zethraeus N, Johnell O, Lidgren L, Ponzer S, Svensson O et al (2006) Costs and quality of life associated with osteoporosis-related fractures in Sweden. Osteoporos Int 17:637–650
111. Peasgood T, Herrmann K, Kanis JA, Brazier JE (2009) An updated systematic review of health state utility values for osteoporosis related conditions. Osteoporos Int 20:853–868
112. Ström O, Borgström F, Zethraeus N, Johnell O, Lidgren L, Ponzer S, Svensson O, Abdon P, Ornstein E, Ceder L, Thorngren KG, Sernbo I, Jonsson B (2008) Long-term cost and effect on quality of life of osteoporosis-related fractures in Sweden. Acta Orthop 79:269–280
113. Parker MJ, Anand JK (1991) What is the true mortality of hip fractures? Public Health 105:443–446
114. Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B (2004) Excess mortality after hospitalisation for vertebral fracture. Osteoporos Int 15:108–112
115. Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B, Oglesby AK (2003) The components of excess mortality after hip fracture. Bone 32:468–473
116. Borgström F, Johnell O, Kanis JA, Jonsson B, Rehnberg C (2006) At what hip fracture risk is it cost-effective to treat? International intervention thresholds for the treatment of osteoporosis. Osteoporos Int 17:1459–1471
117. Zethraeus N, Ström O, Borgström F (2006) What is the risk of institutionalization after hip fracture? Osteoporos Int 17(Suppl 2):60
118. McLellan AR. Reid DM, Forbes K et al (1999) Effectiveness of strategies for the secondary prevention of osteoporotic fractures in Scotland. CEPS99/03. www.nhshealthquality.org/nhsqis/controller?p_service=Content.show&p_applic=CCC&pContentID=2755 Accessed 6th May 2007
119. Kanis JA, Johnell O, Oden A et al (2004) The risk and burden of vertebral fractures in Sweden. Osteoporos Int 15:20–26
120. Kind P, Dolan P, Gudex C, Williams A (1998) Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ 316:736–741
121. Tosteson AN, Melton LJ 3rd, Dawson-Hughes B, National Osteoporosis Foundation Guide Committee et al (2008) Cost-effective osteoporosis treatment thresholds: the United States perspective. Osteoporos Int 19:437–447
122. Brown J, Josse RG, for the Scientific Advisory Council of the Osteoporosis Society of Canada (2002) Clinical practice guidelines for the diagnosis and management of osteoporosis in Canada. Can Med Assoc J 167(suppl 10):S1–S34
123. Kanis JA, Black D, Cooper C, Dargent P, Dawson-Hughes B, De Laet C, on behalf of the International Osteoporosis Foundation and National Osteoporosis Foundation et al (2002) A new approach to the development of assessment guidelines for osteoporosis. Osteoporos Int 13:527–536
124. Borgström F, Carlsson A, Sintonen H, Boonen S, Haentjens P, Burge R et al (2006) The cost-effectiveness of risedronate in the treatment of osteoporosis: an international perspective. Osteoporos Int 17:996–1007
125. Borgström F, Jonsson B, Ström O, Kanis JA (2006) An economic evaluation of strontium ranelate in the treatment of osteoporosis in a Swedish setting: based on the results of the SOTI and TROPOS trials. Osteoporos Int 17:1781–1783
126. Johnell O, Kanis JA (2006) An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int 17:1726–1733
127. Zethraeus N, Ström O, Borgström F, Kanis JA, Jönsson B (2008) The cost-effectiveness of the treatment of high risk women with osteoporosis, hypertension and hyperlipidaemia in Sweden. Osteoporos Int 19:819–828