

Impact of prevalent and incident vertebral fractures on utility: results from a patient-based and a population-based sample

N. M. van Schoor · S. K. Ewing · T. W. O'Neill · M. Lunt · J. H. Smit · P. Lips

Accepted: 12 November 2007
© Springer Science+Business Media B.V. 2007

Abstract Data are scarce on the impact of vertebral fractures (VFX) on utility. The objective of this study was to assess the impact of prevalent and incident VFX on utility in both a patient-based and population-based sample. Data from the Multiple Outcomes of Raloxifene Evaluation (MORE) study ($n = 550$ for prevalent VFX and $n = 174$ for incident VFX) and the European Prospective Osteoporosis Study (EPOS) ($n = 236$) were used. Utility was assessed by the index score of the EQ-5D. In the MORE study, highly statistically significant associations were found between utility and the presence of prevalent VFX ($p < 0.001$), number of prevalent VFX ($p < 0.001$), severity of prevalent VFX ($p < 0.001$), the combination of number and severity of prevalent VFX ($p = 0.001$) and location of prevalent VFX ($p = 0.019$). The mean utility

was significantly lower among women who suffered an incident VFX (utility = 0.67) than among women who did not (utility = 0.77) ($p = 0.005$), although utility loss was not significantly different between the two groups ($p = 0.142$). In EPOS, the combination of number and severity of incident VFX was significantly related to utility ($p = 0.030$). In conclusion, utility is lower among persons with prevalent and incident VFX, especially in a patient-based sample. Utility loss was not significantly different between women without and with incident VFX.

Keywords Incident · Prevalent · Utility · Vertebral fracture

Abbreviations

VFX	Vertebral fractures
MORE study	Multiple Outcomes of Raloxifene Study
EPOS	European Prospective Osteoporosis Study
EVOS	European Vertebral Osteoporosis Study
BMD	Bone mineral density
QALY	Quality-adjusted life years
SQ	Semi-quantitative
BSQ	Binary semi-quantitative
QM	Quantitative morphometric
BMI	Body mass index

N. M. van Schoor (✉) · P. Lips
EMGO Institute, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands
e-mail: nm.vanschoor@vumc.nl

S. K. Ewing
Prevention Sciences Group, UCSF, San Francisco, CA, USA

T. W. O'Neill · M. Lunt
ARC Epidemiology Unit, University of Manchester, Manchester, UK

J. H. Smit
Department of Sociology and Social Gerontology, Vrije Universiteit, Amsterdam, The Netherlands

J. H. Smit
Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands

P. Lips
Department of Endocrinology, VU University Medical Center, Amsterdam, The Netherlands

Introduction

Vertebral fractures (VFX) are among the most common osteoporotic fractures. The prevalence and incidence figures vary strongly between different studies [1–7]. In a large European study, the European Vertebral Osteoporosis

Study (EVOS), the prevalence of radiographically defined VFX was 12% in both women and men [8]. This study was continued as the European Prospective Osteoporosis Study (EPOS), in which the age-standardized incidence of VFX was estimated at 10.7/1,000 person years in women and 5.7/1,000 person years in men [9]. Both clinical and sub-clinical VFX are associated with back pain, impaired physical functioning, and decreased quality of life [1, 3, 4, 7, 10–19].

Quality of life can be assessed by either generic questionnaires, such as the EQ-5D and SF-36, or disease-specific questionnaires, such as the Qualeffo-41 and OPAQ [20]. In addition, some generic questionnaires can be used to calculate utilities. This is the value attached to a specific health state, which varies between 0 (representing death) and 1 (representing perfect health) [20]. Advantages are that multiple outcomes with different consequences, e.g., VFX and hip fracture, can be converted into a single score, and that this single score can be used to compare different diseases. Furthermore, utilities can be used to calculate loss of quality-adjusted life years (QALYs), a measure increasingly used in cost-effectiveness studies [21].

In the field of osteoporosis, there is a lack of empirical data on utility after different types of fractures [21]. Only few studies have assessed the association between VFX and utility [1, 11, 14, 16–19]. The utility after VFX varied between 0.20 and 0.85, depending on the time frame, definition of VFX, and population used. Most studies did not examine incident VFX, or differentiated between severity and location of the VFX. Therefore, the objective of our study was to assess the impact of prevalent and incident VFX on utility in both a patient-based sample (from a clinical trial) and a population-based sample (from a cohort study), and to distinguish between number, severity and location of VFX.

Methods

Patient-based sample

The Multiple Outcomes of Raloxifene Evaluation (MORE) study is a double-blind, placebo-controlled, randomized clinical trial examining the effects of raloxifene in postmenopausal women. Patients were assigned to study group 1 if they had low bone mineral density (BMD) (T -score ≤ -2.5) and no more than one mild prevalent VFX (see Fig. 1). Study group 2 included women with either at least one moderate VFX or at least two mild VFX in the presence of low BMD, or at least two moderate fractures regardless of BMD. To assess the association of prevalent VFX and utility, baseline EQ-5D data were gathered in study group 1 ($n = 304$) and 2 ($n = 346$) in four European

countries: Belgium, Sweden, The Netherlands, and the United Kingdom. To assess the association of incident VFX and utility, only patients from study group 2 ($n = 346$) were included since only they were scheduled to have quality of life data gathered at follow-up visits. Because patients from study group 2 were included in both the prevalent and incident VFX analyses, the study samples were not independent.

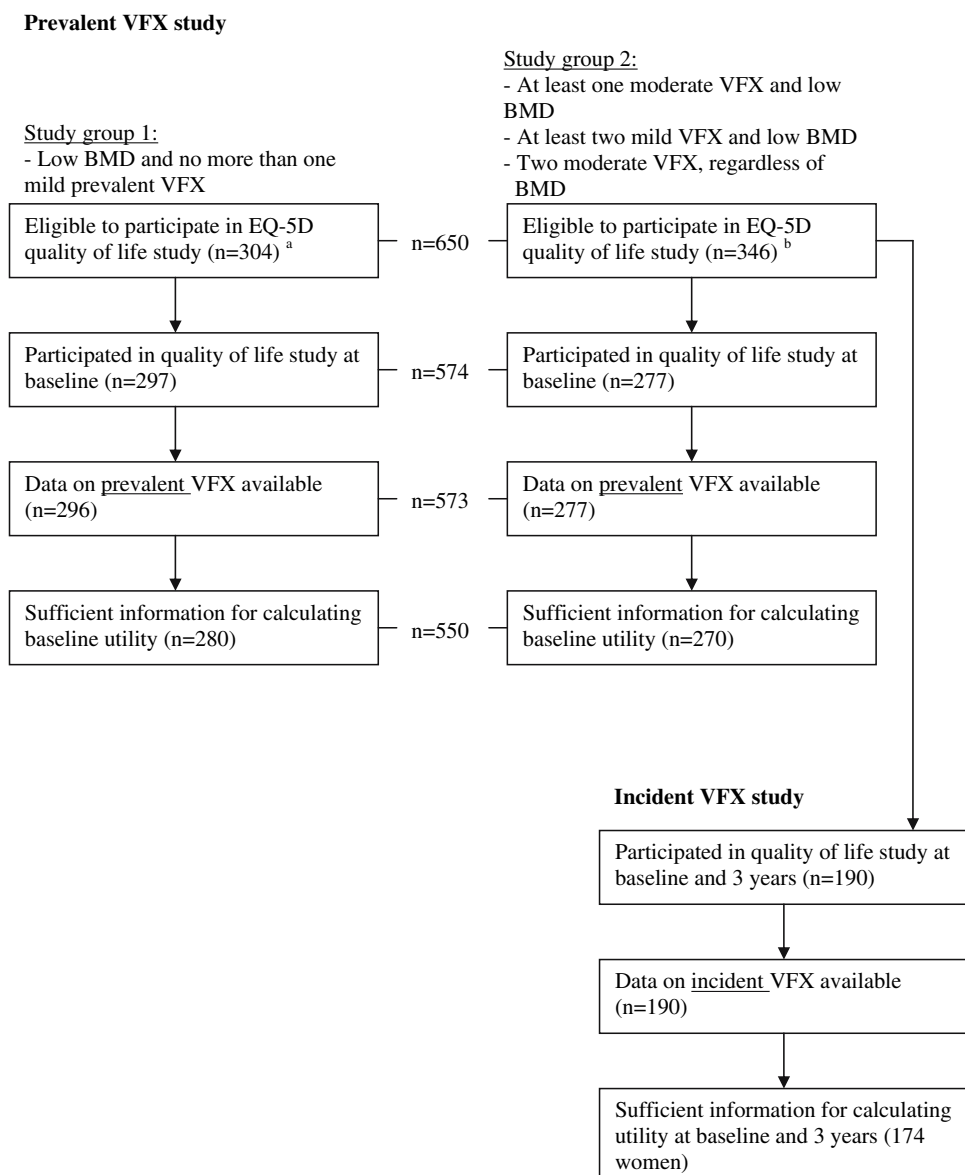
Population-based sample

EPOS is a cohort study, in which subjects from population-based registers in 36 centers were recruited. Stratified random sampling was used with the aim of recruiting a target number of 50 subjects in each of six 5-year age and sex bands per center: 50–54, 55–59, 60–64, 65–69, 70–75, and 75 years and over. Subjects in 29 centers were invited for a baseline and follow-up radiograph of the thoracic and lumbar spine, and 12 centers in seven countries (Czech Republic, Hungary, Italy, Slovakia, Spain, Sweden and United Kingdom) agreed to participate in the quality of life study. The quality of life study is a substudy ($n = 269$), carried out as a nested case-control study, in which first all cases with incident VFX were selected ($n = 73$). It was then attempted to recruit three controls per case, matched for age (within 5 years), sex and center: one with a prevalent deformity at baseline and two without. In practice, only 60 and 136 controls were recruited in these two groups, respectively. Subjects with prevalent VFX were only included if they also had an incident VFX, or they were recruited as a matched control, and hence they are not representative of prevalent VFX in the population. As a consequence, prevalent VFX could not be analyzed separately.

Methods of assessing VFX

In the MORE study, spinal radiographs were taken at baseline, and at 2 and 3 years after baseline. Before randomization, standardized baseline spinal radiographs were assessed quantitatively. Vertebral deformities were assigned to anterior, central or posterior height loss in comparison with the adjacent vertebra. Mild prevalent VFX were described as 20–25% height loss and moderate as 25% height loss or more. Three different methods were used to assess whether an incident VFX occurred during follow-up: the semi-quantitative method (SQ), the binary semi-quantitative method (BSQ), and the quantitative morphometric method (QM) [10, 14]. Radiologists were blinded to treatment assignment but not to the temporal sequence of the radiographs. The SQ method was applied to each participant's baseline and follow-up radiographs using a

Fig. 1 Flow chart MORE study. **a** Study group 1 participants from The Netherlands and United Kingdom were scheduled to participate in the EQ-5D quality of life study. **b** Study group 2 participants from Belgium, Sweden, The Netherlands and United Kingdom were scheduled to participate in the EQ-5D quality of life study



score of 0 (no fracture), 1.0 (mild), 2.0 (moderate), or 3.0 (severe). A change in SQ score of 1.0 from baseline to follow-up resulted in BSQ and QM analyses performed on the entire series of radiographs for the participant. The BSQ analysis was performed by a second radiologist who rated the fracture as 0 (no fracture) or 1.0 (fracture). For QM, both a 20% height decrease and a 4-mm absolute vertebral height decrease was required for an incident fracture to be diagnosed. An incident VFX was reported only when confirmed by at least two of the three readings. In EPOS, both baseline and follow-up spinal radiographs (mean = 3.8 years after baseline) were evaluated in a single center (Berlin) [12, 22]. Prevalent deformities at baseline were defined morphometrically using the McCloskey–Kanis algorithm [23]. Incident VFX were classified both by qualitative (clinical) radiologist assessment and/or

morphometrically. The morphometric method used required a vertebra to satisfy criteria for a McCloskey–Kanis deformity on the second radiograph and, in addition, show a change (between films) in either anterior, mid or posterior vertebral height of at least 20% (with the reduction in height being at least 4 mm).

Utility

Utility was assessed by the EQ-5D in both studies [24]. For the MORE study, baseline utility and utility at 3 years after baseline were used. In EPOS, utility was assessed only once, at 1.9 years (= median) after the second follow-up (approximately 3.8 + 1.9 = 5.7 years after baseline). The

EQ-5D consists of five items representing the domains mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These can be converted into a single index score or utility, which summarizes the health state of a participant. In total, there are 243 different health states possible. In a UK study, the time trade-off method was used to derive valuations for 42 health states, which were subsequently modeled to predict the remaining index scores [25].

Statistical analyses

The analyses of prevalent and incident VFX, respectively, with utility were largely performed in the same way. First, mean utilities and 95% confidence intervals were calculated for persons without and with VFX. To test whether the groups were significantly different, Mann–Whitney *U*-test and Kruskal–Wallis *H*-test were used because of the skewed distribution of utility. In the MORE study, the influence of number, severity, a combination of number and severity, and location of VFX on utility was examined only for prevalent VFX, since power was not sufficient to analyze these characteristics for incident VFX. In EPOS, the influence of number, severity and location of VFX was only examined for incident VFX, while the data were not available for prevalent VFX.

Because of several baseline differences between women with and without prevalent VFX in the MORE study, presence of prevalent VFX was also stratified by median age, median body mass index (BMI) and history of non-VFX. Due to power limitations, we did not stratify the incident VFX analyses.

In the MORE study, change in utility was calculated by subtracting baseline utility from utility at 3 years in order to compare loss of utility in women with and without an incident VFX during follow-up.

The combination of number and severity of VFX in both studies was calculated as follows: all persons with severe VFX, all persons with three or more VFX and all persons with two VFX of which at least one was moderate were categorized as “severe osteoporosis”, and the remainder as “mild or no osteoporosis”.

Results

For the prevalent VFX analyses of the MORE study, data on quality of life were available for 574 women, and data on prevalent VFX were available for 573 of these women (see Fig. 1). Baseline utilities could be calculated for 550 women. Non-responders (650–550 = 100 women) were significantly older ($p = 0.01$) and had a higher number of

postmenopausal years ($p < 0.001$) than responders. For the incident VFX analyses, data on quality of life and incident VFX were available for 190 women (see Fig. 1). Baseline and 3-year utilities could be calculated for 174 women. Non-responders (346–174 = 172 persons) were significantly older ($p < 0.001$) and had a higher number of postmenopausal years ($p < 0.001$) than responders. In EPOS, utilities could be calculated for 63 cases and 173 controls. Non-responders (269–236 = 33 persons) had significantly more fractures before baseline than responders ($p = 0.032$).

In Tables 1 and 2, the baseline characteristics of the MORE study and EPOS are presented. In the MORE study, women with prevalent VFX were significantly older, had a higher BMI, more years since menopause, a lower lumbar spine BMD, and a higher percentage with a history of non-VFX than women without prevalent VFX. Furthermore, women with incident VFX had significantly lower lumbar spine BMD than women without incident VFX. In EPOS, no significant baseline differences were found.

In Table 3, it can be seen that the presence of prevalent VFX ($p < 0.001$), an increasing number of prevalent VFX ($p < 0.001$), severity of prevalent VFX ($p < 0.001$), and the combination of number and severity of prevalent VFX ($p = 0.001$) were highly significantly associated with utility in the MORE study. With regard to the location of the prevalent VFX, the utility scores were significantly different between women who had only a thoracic VFX, women who had only a lumbar VFX, and women who had both a thoracic and lumbar VFX ($p = 0.019$). However, when dividing the thoracic and lumbar regions into more precise locations, no significant differences were found between the locations ($p = 0.853$).

In Table 4 the influence of age, BMI and history of non-VFX on the association between utility and prevalent VFX is presented. Women with prevalent VFX consistently had significantly lower baseline utility scores compared to women without prevalent VFX. The adverse effect of prevalent VFX on utility was most pronounced among women having a high BMI.

When comparing all women without incident VFX ($n = 143$) during the 3 years of follow-up of the MORE study to those women who did have an incident VFX ($n = 31$), a statistically significant difference was found at baseline, thus before the incident VFX occurred [utility for no incident VFX = 0.76 (95%CI: 0.72–0.80); utility for incident VFX = 0.69 (95%CI: 0.60–0.77); $p = 0.021$]. In addition, a highly significant difference after 3 years was observed [utility for no incident VFX = 0.77 (95%CI: 0.73–0.81); utility for incident VFX = 0.67 (95%CI: 0.58–0.76); $p = 0.005$]. However, utility loss during the 3-year follow-up period was not significantly different between women without and with incident VFX ($p = 0.142$).

Table 1 Baseline characteristics of the MORE study

	Prevalent VFX			Incident VFX ^a		
	No (<i>n</i> = 271)	Yes (<i>n</i> = 279)	<i>p</i> -value ^b	No (<i>n</i> = 143)	Yes (<i>n</i> = 31)	<i>p</i> -value ^b
Age (years) ^c	66.2 ± 5.9	68.8 ± 6.3	<0.001	67.7 ± 6.0	68.8 ± 6.2	0.377
BMI (kg/cm ²) ^c	24.9 ± 3.4	25.6 ± 3.8	0.028	25.5 ± 3.6	25.7 ± 4.1	0.706
Years postmenopausal ^c	17.8 ± 7.1	21.0 ± 7.8	<0.001	19.8 ± 8.0	20.9 ± 8.1	0.482
Caucasian (%)	267 (98.5%)	275 (98.6%)	1.00	141 (98.6%)	30 (96.8%)	0.479
Lumbar spine BMD (g/cm ²) ^c	0.83 ± 0.14	0.78 ± 0.13	<0.001	0.79 ± 0.13	0.72 ± 0.09	0.010
History of non-VFX (%)	86 (31.7%)	131 (47.0%)	<0.001	76 (53.1%)	17 (54.8%)	0.864

^a By design, all persons selected for the follow-up study had prevalent vertebral fractures

^b Differences in mean were tested with *T*-test, and in proportion with χ^2 test

^c Mean ± SD

Table 2 Baseline characteristics of EPOS^a

	Controls without incident VFX			Cases with incident VFX			<i>p</i> -value ^b
	No prevalent VFX (<i>n</i> = 120)	Prevalent VFX (<i>n</i> = 53)	All controls (<i>n</i> = 173)	No prevalent VFX (<i>n</i> = 43)	Prevalent VFX (<i>n</i> = 20)	All cases (<i>n</i> = 63)	
Age (years) ^c	64.4 ± 6.8	63.4 ± 7.1	64.1 ± 6.8	63.9 ± 7.5	67.8 ± 6.0	65.2 ± 7.2	0.301
Sex (% female)	95 (79.2%)	39 (73.6%)	134 (77.5%)	35 (81.4%)	14 (70.0%)	49 (77.8%)	0.958
BMI (kg/cm ²) ^c	27.3 ± 4.1	26.5 ± 3.6	27.1 ± 4.0	27.3 ± 3.7	26.0 ± 3.5	26.9 ± 3.7	0.789
Lumbar spine BMD ^{d,e}	0.94 (0.88–1.15)	0.97 (0.79–1.09)	0.95 (0.88–1.11)	0.91 (0.74–1.00)	0.83 (0.68–0.96)	0.89 (0.72–0.98)	0.079
Pre-baseline fracture since age 50 (%)	14 (11.7%)	10 (18.9%)	24 (13.9%)	6 (14.0%)	6 (30.0%)	12 (19.0%)	0.328

^a While the EPOS study is a case-control study in which controls were selected to match the cases (with incident VFX) by age, sex and center, the controls are not representative for the general population. Per case, two controls were selected who did not have a previous vertebral fracture and one control who did have a previous vertebral fracture

^b Comparison: all controls versus all cases with an incident vertebral fracture. Differences in mean were tested with *T*-test, in median with Mann–Whitney *U* test, and in proportion with χ^2 test

^c Mean ± SD

^d Median (and interquartile range)

^e Only assessed in a subgroup (*n* = 72)

In EPOS (Table 5), a significantly lower utility was observed when combining number and severity of VFX (*p* = 0.030). No other significant differences were observed.

Discussion

A lower utility was observed among women with prevalent VFX as compared with women having no VFX in a patient-based sample. Although utility was significantly lower among women who suffered an incident VFX, utility loss was not significantly different. This indicates that the occurrence of incident VFX can already be predicted at an early stage by a worse health state. In a population-based study, only the combination of number and severity of incident VFX was significantly related to utility.

In both studies, the lowest utility was found for persons with incident VFX who already had prevalent VFX. All persons in our analysis from the MORE study with incident VFX also had prevalent VFX, since follow-up EQ-5D data were collected only in women in study group 2 (utility = 0.67). In EPOS, the utility was 0.71 for cases with incident VFX who also had prevalent VFX. These findings are consistent with those reported in an earlier EPOS study which showed that cases with a baseline deformity had significantly lower utility than controls without a baseline deformity (*p* < 0.001) [12]. Furthermore, in the MORE study, especially a high BMI in combination with prevalent VFX had a great impact on utility. An earlier EPOS study showed that women with lumbar deformity were more likely to report back pain than those with a thoracic deformity [26]. In our study, utility was lower in women

Table 3 Association between prevalent vertebral fractures and mean utility (95% CI) in the MORE study

Prevalent VFX	Number of persons	Utility at baseline	<i>p</i> -value ^a
Presence of prevalent VFX			
No	271	0.82 (0.80–0.85)	<0.001
Yes	279	0.73 (0.70–0.76)	
Number of prevalent VFX ^b			
0	271	0.82 (0.80–0.85)	<0.001
1	122	0.75 (0.71–0.79)	
2	65	0.74 (0.68–0.80)	
≥3	92	0.71 (0.66–0.77)	
Most severe prevalent VFX ^c			
No	253	0.82 (0.80–0.85)	<0.001
Mild	140	0.77 (0.74–0.80)	
Moderate	89	0.70 (0.63–0.76)	
Severe	68	0.72 (0.66–0.78)	
Combination of number and severity of prevalent VFX			
Mild or no osteoporosis	410	0.80 (0.78–0.82)	0.001
Severe osteoporosis	140	0.72 (0.67–0.76)	
Location of prevalent VFX (1) ^b			
Thoracic vfx only	136	0.78 (0.74–0.81)	0.019 ^d
Lumbar vfx only	41	0.68 (0.57–0.79)	
Thoracic and lumbar VFX	102	0.70 (0.65–0.75)	
Location of prevalent VFX (2) ^e			
T4–T6 only	13	0.77 (0.65–0.90)	0.853 ^c
T7–T9 only	45	0.79 (0.74–0.84)	
T10–T12 only	51	0.76 (0.70–0.82)	
L1–L2 only	30	0.72 (0.60–0.83)	
L3–L4 only	6	0.76 (0.62–0.90)	

^a Differences in means were tested using Mann–Whitney *U*- and Kruskal–Wallis *H*-tests

^b Mean utility in women with 0, 1 and 2 VFX, thoracic VFX only and lumbar VFX only were also reported in Oleksik et al. [14]

^c As assessed by the SQ method which uses a score of 0 (no fracture), 1.0 (mild), 2.0 (moderate), and 3.0 (severe)

^d *p*-value represents comparison among women *with* prevalent VFX. When adding the women without VFX to these analyses ($n = 271$; utility = 0.82; 95% CI: 0.80–0.85), the difference between the groups is statistically significant at $p < 0.001$

^e The number of persons do not add up to the total number of incident VFX because persons who had incident VFX at two different locations were not included in this analysis. When adding the women with multiple locations to the analysis ($n = 134$; utility = 0.70; 95% CI: 0.65–0.75), the difference between the groups is statistically significant at $p < 0.001$

with lumbar VFX, but this difference was not statistically significant.

Most associations between VFX and utility were only statistically significant in the MORE study, and not in EPOS. This may be explained by several reasons. First, the MORE study is a patient-based sample in which the patients were selected in a clinic. In contrast, the participants from EPOS, although a relatively small case-control study which might not be completely representative of the general population anymore, were selected in the population. Therefore, it is likely that more subclinical (asymptomatic) VFX were included. Second, the time-frame between incident VFX and assessment of the EQ-5D instrument is longer in EPOS than in the MORE study. In the MORE study, the follow-up radiographs were at 2 and

3 years after baseline; in EPOS, the follow-up radiograph was at a mean of 3.8 years after baseline. Furthermore, in the MORE study, quality of life was assessed at 3 years and in EPOS at a median of 1.9 years after the second follow-up radiograph. Assuming that the incident VFX occurred in the middle between the two radiographs, this means that utility was assessed 1.5 years after the incident VFX in the MORE study, and 3.8 years after the incident VFX in EPOS. Finally, the sample size was larger in the MORE study and therefore the MORE study had more power to detect statistically significant differences.

The most important strength of this study is that it differentiates between number, severity and location of VFX, and that it not only examined prevalent VFX, but also incident VFX. Therefore, this study contributes to the

Table 4 Influence of age, body mass index and history of non-vertebral fractures on the association between prevalent vertebral fractures and mean utility (95% CI) in the MORE study

Stratum	Prevalent VFX	Number of persons	Utility at baseline	<i>p</i> -value ^a
Age				
<67.8	No	159	0.83 (0.79–0.86)	<0.001
	Yes	116	0.73 (0.69–0.78)	
≥67.8	No	112	0.81 (0.78–0.85)	0.018
	Yes	163	0.73 (0.69–0.77)	
BMI				
<25.1	No	145	0.83 (0.79–0.86)	0.007
	Yes	130	0.78 (0.74–0.81)	
≥25.1	No	126	0.82 (0.78–0.85)	<0.001
	Yes	148	0.70 (0.65–0.74)	
History of non-VFX				
No	No	185	0.83 (0.80–0.86)	<0.001
	Yes	148	0.75 (0.70–0.78)	
Yes	No	86	0.80 (0.76–0.85)	0.014
	Yes	131	0.72 (0.68–0.76)	

^a Differences in means were tested using Mann–Whitney *U*-test

knowledge on utility in persons with osteoporosis. In addition, both studies were prospective studies.

A limitation might be that it was not possible to normalize the distribution of utility due to multiple peaks in order to perform adjusted analyses. However, utilities are

usually presented unadjusted. Because quality of life is such a broad concept, many factors are part of it and it may be too strict to adjust for these factors. Baseline differences, such as BMI, may influence mobility, for example, and therefore quality of life. However, we did stratify the analyses on age, BMI and history of non-VFX in order to examine the influence of these factors on utility.

In the MORE study, only small differences were observed between one and two prevalent VFX, moderate and severe prevalent VFX, lumbar VFX and both thoracic and lumbar VFX (Table 3). Furthermore, there was little utility loss in women having incident VFX. In EPOS, no significant differences were observed for incident VFX (yes/no), number of incident VFX, most severe incident VFX and location of VFX (Table 5). These results indicate that the EQ-5D might be less sensitive to change than the more specific quality of life questionnaires, such as Qual-effo-41. A disadvantage of the EQ-5D is that it consists of only five questions, which may make the instrument less valid for describing the impact of certain diseases. Another problem might be that only 42 out of 243 health states of the EQ-5D were derived using the time trade-off method, and the remaining ones were predicted by modeling these 42 health states. However, the results for the MORE study might also be explained by a threshold in the impact of number and severity of VFX on utility, and by the fact that lumbar VFX are known to have a greater impact than thoracic VFX. Furthermore, in prevalent VFX it is not clear

Table 5 Association between incident vertebral fractures and mean utility (95% CI) in EPOS

Incident VFX	Number of persons	Utility 5.7 years after baseline	<i>p</i> -value ^a
Controls without incident VFX			
No prevalent VFX	120	0.82 (0.79–0.85)	0.065 ^b
Prevalent VFX	53	0.80 (0.75–0.85)	
All controls	173	0.82 (0.79–0.84)	
Cases with incident VFX			
No prevalent VFX	43	0.80 (0.73–0.86)	
Prevalent VFX	20	0.71 (0.63–0.78)	
All cases	63	0.77 (0.72–0.82)	
Number of incident VFX ^c			
0	187	0.81 (0.78–0.84)	0.406
1	40	0.78 (0.72–0.83)	
≥2	9	0.79 (0.68–0.91)	
Most severe incident VFX ^c			
No (<20% height loss)	187	0.81 (0.78–0.84)	0.200
Mild (20–25% height loss)	7	0.77 (0.55–1.00)	
Moderate (25–40% height loss)	25	0.82 (0.77–0.87)	
Severe (>40% height loss)	17	0.71 (0.62–0.81)	
Combination of number and severity of incident VFX ^c			
Mild or no osteoporosis	215	0.81 (0.79–0.83)	0.030

Table 5 continued

Incident VFX	Number of persons	Utility 5.7 years after baseline	<i>p</i> -value ^a
Severe osteoporosis	21	0.73 (0.65–0.81)	
Location of incident VFX ^{c,d}			
Thoracic vfx only	28	0.80 (0.74–0.87)	0.447 ^e
Lumbar vfx only	18	0.75 (0.66–0.84)	

^a Differences were tested using Mann–Whitney *U*- and Kruskal–Wallis *H*-tests

^b Comparison: all controls versus all cases with an incident VFX. Comparisons of all cases, cases without baseline deformity and cases with baseline deformity versus controls without baseline deformity and controls with baseline deformity were reported in Cockerill et al. [12] and are not reported here

^c Morphometric fractures

^d The number of persons do not add up to the total number of incident VFX because persons who had incident VFX at two different locations were not included in this analysis

^e *p*-value represents comparison of persons *with* incident VFX. When adding the persons without VFX to the analysis ($n = 187$; utility = 0.81; 95% CI: 0.78–0.84), the *p*-value = 0.347

how long the VFX is present, and this may influence the impact on utility. The lack of significant effects as observed in EPOS might be explained by low numbers and by the inclusion of asymptomatic fractures. Finally, in both studies, part of the incident VFX might have occurred some time ago because of the time period between the radiographs. Future research should confirm the sensitivity and validity of the EQ-5D instrument.

Other limitations were that, for some categories of VFX, the number of persons was too small to assess the association between these categories and utility, and that multiple tests were carried out increasing the risk of type I error. Furthermore, we were not able to examine different time frames as for example in the studies of Johnell et al. and Tosteson et al. [17, 18]. In the first study, part of the lost quality of life was regained after 6 months, but after that a stable reduction was found until 12 months. In the second study, utility after VFX changed from 0.74 within 1 year after the fracture to 0.85 after more than 2 years. It would be interesting to further examine this in future prospective studies with larger numbers of VFX and shorter intervals between the VFX and utility assessment. Finally, particularly in the MORE study, it was not possible to calculate utilities for all participants. Since non-responders were significantly older, the associations presented here may have been underestimated.

In conclusion, utility as assessed by the index score of the EQ-5D is lower among persons with prevalent and incident VFX, especially in a patient-based sample. Utility loss was not significantly different between women without and with incident VFX. The findings of this study may contribute to the knowledge on the impact of VFX on utility in the field of osteoporosis.

Acknowledgements Natasja van Schoor was funded by an unconditional grant from Wyeth Research, Collegeville, Pennsylvania. Furthermore, we would like to acknowledge the investigators and

participants of the MORE study and EPOS. MORE study: Eli Lilly and Company provided funding for the Multiple Outcomes of Raloxifene Evaluation Trial. These data were presented in part at the Fifth European Congress on Clinical and Economical Aspects of Osteoporosis and Osteoarthritis in 2005. The complete list of all investigators in the MORE study was previously published in: Ettinger B, Black DM et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. *JAMA* 1999; 282, 637–645. Epos: Principal Investigators: Prague (Czech Republic) J Stepan; Berlin (Germany) D Felsenberg; Lubeck (Germany) H Raspe; Budapest (Hungary) C Kiss, G Poor; Piastany (Slovakia) P Masaryk; Oviedo (Spain) JB Cannata, M Naves; Malmo (Sweden) O Johnell; Aberdeen (UK) DM Reid; Bath (UK) A Bhalla; Cambridge (UK) C Todd; Harrow (UK) J Reeve; Sheffield (UK) R Eastell; Truro (UK) AD Woolf. EPOS was financially supported by a European Union Concerted Action Grants under Biomed-1 (BMH 1CT 931448 and 920182), and also EU grants C1PDCT925102, ERBC1PDCT 930105 and 940229. The central coordination was also supported by the UK Arthritis Research Campaign, the Medical Research Council (G9321536), and the European Foundation for Osteoporosis and Bone Disease. Also supported by a grant from Eli Lilly and Company. The EU's PECO program linked to BIOMED 1 funded in part the participation of the Budapest, Prague and Piastany centers. Data collection in Cambridge was supported by a HSR grant from the Anglia and Oxford region. The central X-ray evaluation was sponsored by the Bundesministerium fur Forschung and Technologie, Germany.

References

1. Kanis, J. A., Johnell, O., Oden, A., Borgstrom, F., Zethraeus, N., de Laet, C., et al. (2004). The risk and burden of vertebral fractures in Sweden. *Osteoporosis International*, 15, 20–26.
2. Hasserijs, R., Redlund-Johnell, I., Mellstrom, D., Johansson, C., Nilsson, B. E., & Johnell, O. (2001). Vertebral deformation in urban Swedish men and women – Prevalence based on 797 subjects. *Acta Orthopaedica Scandinavica*, 72, 273–278.
3. Pluijm, S. M., Tromp, A. M., Smit, J. H., Deeg, D. J., & Lips, P. (2000). Consequences of vertebral deformities in older men and women. *Journal of Bone and Mineral Research*, 15, 1564–1572.
4. Lau, E. M. C., Woo, J., Chan, H., Chan, M. K. F., Griffith, J., Chan, Y. H., et al. (1998). The health consequences of vertebral deformity in elderly Chinese men and women. *Calcified Tissue International*, 63, 1–4.

5. Jones, G., White, C., Nguyen, T., Sambrook, P. N., Kelly, P. J., & Eisman, J. A. (1996). Prevalent vertebral deformities: Relationship to bone mineral density and spinal osteophytosis in elderly men and women. *Osteoporosis International*, *6*, 233–239.
6. Melton, L. J. III, Lane, A. W., Cooper, C., Eastell, R., O’Fallon, W. M., & Riggs, B. L. (1993). Prevalence and incidence of vertebral deformities. *Osteoporosis International*, *3*, 113–119.
7. Silverman, S. L. (1992). The clinical consequences of vertebral compression fracture. *Bone*, *13*(Suppl 2), S27–S31.
8. O’Neill, T. W., Felsenberg, D., Varlow, J., Cooper, C., Kanis, J. A., Silman, A. J., et al. (1996). The prevalence of vertebral deformity in European men and women: The European Vertebral Osteoporosis Study. *Journal of Bone and Mineral Research*, *11*, 1010–1018.
9. Felsenberg, D., Silman, A. J., Lunt, M., Armbrrecht, G., Ismail, A. A., Finn, J. D., et al. (2002). Incidence of vertebral fracture in Europe: Results from the European Prospective Osteoporosis Study (EPOS). *Journal of Bone and Mineral Research*, *17*, 716–724.
10. Oleksik, A. M., Ewing, S., Shen, W., van Schoor, N. M., & Lips, P. (2005). Impact of incident vertebral fractures on health related quality of life (HRQOL) in postmenopausal women with prevalent vertebral fractures. *Osteoporosis International*, *16*, 861–870.
11. van Schoor, N. M., Smit, J. H., Twisk, J. W. R., & Lips, P. (2005). Impact of vertebral deformities, osteoarthritis, and other chronic diseases on quality of life: A population-based study. *Osteoporosis International*, *16*, 749–756.
12. Cockerill, W., Lunt, M., Silman, A. J., Cooper, C., Lips, P., Bhalla, A. K., et al. (2004). Health-related quality of life and radiographic vertebral fracture. *Osteoporosis International*, *15*, 113–119.
13. Adachi, J. D., Loannidis, G., Berger, C., Joseph, L., Papaioannou, A., Pickard, L., et al. (2001). The influence of osteoporotic fractures on health-related quality of life in community-dwelling men and women across Canada. *Osteoporosis International*, *12*, 903–908.
14. Oleksik, A., Lips, P., Dawson, A., Minshall, M. E., Shen, W., Cooper, C., et al. (2000). Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *Journal of Bone and Mineral Research*, *15*, 1384–1392.
15. Lips, P., Cooper, C., Agnusdei, D., Caulin, F., Egger, P., Johnell, O., et al. (1999). Quality of life in patients with vertebral fractures. Validation of the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO). *Osteoporosis International*, *10*, 150–160.
16. Hall, S. E., Criddle, R. A., Comito, T. L., & Prince, R. L. (1999). A case–control study of quality of life and functional impairment in women with long-standing vertebral osteoporotic fracture. *Osteoporosis International*, *9*, 508–515.
17. Johnell, O., Kanis, J., deLaet, C., Jonsson, B., Zethraeus, N., & Oden, A. (2002). Sequential changes in quality of life after osteoporotic fractures. *Osteoporosis International*, *13*, S70.
18. Tosteson, A. N., Gabriel, S. E., Grove, M. R., Moncur, M. M., Kneeland, T. S., & Melton, L. J. (2001). Impact of hip and vertebral fractures on quality-adjusted life years. *Osteoporosis International*, *12*, 1042–1049.
19. Gabriel, S. E., Kneeland, T. S., Melton, L. J., Moncur, M. M., Ettinger, B., & Tosteson, A. N. (1999). Health-related quality of life in economic evaluations for osteoporosis: Whose values should we use? *Medical Decision Making*, *19*, 141–148.
20. Lips, P., & Van Schoor, N. M. (2005). Quality of life in patients with osteoporosis. *Osteoporosis International*, *16*, 749–756.
21. Brazier, J. E., Green, C., & Kanis, J. A. (2002). A systematic review of health state utility values for osteoporosis-related conditions. *Osteoporosis International*, *13*, 768–776.
22. Lunt, M., Ismail, A. A., Felsenberg, D., Cooper, C., Kanis, J. A., Reeve, J., et al. (2002). Defining incident vertebral deformities in population studies: A comparison of morphometric criteria. *Osteoporosis International*, *13*, 809–815.
23. McCloskey, E. V., Spector, T. D., Eyres, K. S., Fern, E. D., O’Rourke, N., Vasikaran, S., et al. (1993). The assessment of vertebral deformity – A method for use in population studies and clinical-trials. *Osteoporosis International*, *3*, 138–147.
24. Brooks, R., Rabin, R., & de Charro, F. (Eds.) (2003). *The measurement and valuation of health status using EQ-5D: A European perspective*. Kluwer Academic Publishers.
25. Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care*, *35*, 1095–1108.
26. Cockerill, W., Ismail, A. A., Cooper, C., Matthis, C., Raspe, H., Silman, A. J., et al. (2000). Does location of vertebral deformity within the spine influence back pain and disability? *Annals of the Rheumatic Diseases*, *59*, 368–371.