

## Comparison of methods for the visual identification of prevalent vertebral fracture in osteoporosis

G. Jiang · R. Eastell · N. A. Barrington · L. Ferrar

Received: 19 August 2003 / Accepted: 8 March 2004 / Published online: 8 April 2004  
© International Osteoporosis Foundation and National Osteoporosis Foundation 2004

**Abstract** The identification of vertebral fracture in osteoporosis is based mainly on the identification of abnormal variation in vertebral shape, but this can be misleading in the presence of a non-fracture deformity or normal variant of vertebral shape. Qualitative identification of vertebral fracture (Qual) is influenced by the subjectivity of the approach, and although more objective, the semiquantitative method (SQ) can be difficult to apply. In addition, there has been little independent evaluation of SQ in relation to other approaches. We aimed to evaluate a new algorithm-based approach for the qualitative identification of vertebral fracture (ABQ) and to compare it with SQ and Qual. Two radiologists reported spinal radiographs for 372 postmenopausal women using Qual (reader 1), and SQ and ABQ (reader 2). Non-fracture deformities and normal variants were also reported using Qual and ABQ. The prevalence of vertebral fracture by subjects was higher for SQ (24%) than for Qual (11%) and ABQ (7%). Agreement was poor between SQ and the other methods, and moderate between Qual and ABQ. Twenty-two women with vertebral fracture were agreed by all three methods, similar to the total identified by ABQ (25 women). Seventeen women diagnosed with fracture by Qual, had non-fracture deformity or normal variant (but no fracture) according to ABQ. Of the women with SQ fractures, 53% and 70% were identified negative for fracture but positive for non-fracture deformity or normal variant by ABQ and Qual. The main sources of discrepancy between SQ and the other methods were Scheuermann's

disease, normal variation, and degenerative change accompanied by short anterior vertebral height. For all methods, bone mineral density (BMD) and BMD Z-scores were lower in women with vertebral fractures than in those with no fractures. Bone mineral density and BMD Z-scores were also lower at the lumbar spine and total body in women with vertebral fractures according to Qual and ABQ than they were for SQ, and were lower in women with SQ fractures agreed by Qual and ABQ, compared with those diagnosed negative for fracture by Qual and ABQ ( $p < 0.01$ ). We conclude that poor agreement between methods arises mainly from difficulties in differentiating true fracture from non-fracture deformity. Our new approach attempts to address this problem but requires further testing in a larger study population.

**Keywords** Osteoporosis · Prevalence · Vertebral fracture

### Introduction

The definition of osteoporotic vertebral fracture in qualitative visual diagnosis is based on individual interpretation of the radiographic appearances. Quantitative morphometry offers a more objective and reproducible approach [1, 2, 3, 4, 5, 6, 7, 8, 9] but has other limitations (for example, differences in the prevalence of vertebral deformity for different quantitative methods, false positive identification of non-fracture deformities and underdiagnosis of mild endplate fractures). The semiquantitative approach (SQ) was developed in an attempt to reduce the subjectivity associated with visual diagnosis of vertebral fracture [10]. The approach is widely used in clinical trials and other studies, but there has been little independent evaluation of SQ in comparison to other methods for the identification of vertebral fractures (most analyses have been performed by the group that originally developed the

G. Jiang · R. Eastell (✉) · L. Ferrar  
Bone Metabolism Group, Section of Human Metabolism,  
Division of Clinical Sciences,  
University of Sheffield, Sheffield, UK  
E-mail: r.eastell@sheffield.ac.uk  
Tel.: +44-114-2714705  
Fax: +44-114-2618775

N.A. Barrington  
Department of Diagnostic Imaging,  
Northern General Hospital,  
Sheffield, UK

SQ method). The SQ method can be difficult to apply accurately for the following reasons. Firstly, visual estimation of reduction in vertebral height or area is difficult to achieve accurately without the aid of direct measurements, and the criterion of 20 to 25% for identification of a grade 1 fracture is ambiguous. Secondly, the expected appearance of a normal vertebra in SQ is based on rectangular shape, with no allowance made for variation in vertebral dimensions at different vertebral levels. Finally, the identification of fracture is based on an apparent reduction in vertebral height or area and radiological characteristics of abnormal change in the vertebral endplate. However, the criteria for abnormal appearances in the endplate are not clearly defined. It is difficult to diagnose a grade 1 (mild) fracture, because there may be typical features of endplate change, yet a reduction in vertebral height of less than 20%. Alternatively, there may be anterior height reduction greater than 20% in relation to the posterior height, but no evidence of change in the endplate.

While it is generally acknowledged that intraobserver and interobserver agreement is poor for qualitative visual identification of vertebral fracture, there has been little investigation into the causes of this. Approaches to the identification of vertebral fracture are currently based on the identification of abnormal variation in vertebral shape, or more specifically in vertebral body height. This is problematic, because such deformations may also be due to normal developmental variants of vertebral shape or non-fracture deformities. These can be difficult to differentiate from true fracture (particularly in cases of mild deformation). In addition, the identification of abnormal vertebral shape can be confounded by obliquity of the vertebral bodies due to poor radiographic technique, or to scoliosis.

The advantage of qualitative visual diagnosis is that the expert reader may rule out abnormal appearances that are unrelated to osteoporotic fracture. However, because it is a subjective method, there may often be disagreement between readers about whether a vertebra is fractured [1, 11, 12]. A systematic approach based on the use of predefined criteria for qualitative identification of vertebral fracture could reduce the subjectivity and improve interobserver agreement. We have formulated an algorithm-based qualitative approach (ABQ) that focuses primarily on the identification of change in the vertebral endplate (rather than reduction in vertebral body height) and includes clearly defined guidelines for the differentiation of vertebral fracture and non-fracture deformity.

The aims of this study were to (1) compare semi-quantitative and qualitative radiological assessment with a new algorithm-based qualitative approach for the identification of vertebral fracture in osteoporosis, (2) identify and evaluate discrepancies between methods, and (3) examine the association between bone mineral density (BMD) and vertebral fractures identified by each of the three approaches.

## Methods

### Participants

The study population comprised 375 women ages 50 to 85 years (mean 64.4 years,  $SD = 9.1$ ) drawn as a random sample from three general practice populations in Sheffield, UK (1990 to 1991) to participate in a study of bone density in women as part of the European Vertebral Osteoporosis Study [13]. Women were excluded if they were infirm, or unable to give informed consent. The response rate was 55%. Full details of the study population have been published elsewhere [14]. Spinal radiographs, bone density scans, and anthropometric data were acquired at baseline (on the same day), and at subsequent follow-up visits. This paper reports a retrospective analysis of radiographs and measurements acquired at the baseline visit.

### Radiography, bone density scans, and anthropometric measurements

Radiographs of the thoracolumbar spine were acquired (anteroposterior and lateral projections) according to a standardized protocol. The focus-to-film distance was 100 cm. Thoracic radiographs were centered over vertebra T7 and the lumbar radiographs over vertebra L3. Bone mineral density (BMD) at the lumbar spine, hip (femur neck, trochanter, and Ward's triangle), and total body was measured by dual-energy X-ray absorptiometry (DXA) using a Lunar DPX densitometer (Lunar, Madison, WI, USA). Baseline data was also collected for age, height (stature), and weight.

### Visual diagnosis of vertebral fracture

Three visual approaches were applied for the identification of vertebral fractures and deformities. Two experienced radiologists reported the radiographs independently. Reader 1 (N.A.B.) specializes in skeletal radiology, and reader 2 (G.J.) is a research radiologist. For Qual and ABQ, vertebral fractures and non-fracture deformities or normal variants were reported separately, and subjects with both vertebral fracture and non-fracture deformity were classified into the fracture group. The three methods were applied as follows:

1. Qualitative visual assessment (Qual). Reader 1 assessed the radiographs qualitatively with the aid of a radiological atlas of normal variants. This is the approach commonly adopted in clinical practice in the UK.
2. Semiquantitative assessment (SQ). Reader 2 applied the semiquantitative approach developed by Genant and colleagues [10]. Vertebral fracture was identified by visual estimation of apparent percentage reduction

in vertebral height greater than 20 to 25% (no direct measurements were performed). The severity of fracture was determined by vertebral height reductions of 20 to 25% (grade 1), 25 to 40% (grade 2), and >40% (grade 3). Minor reductions in vertebral height of <20%, or equivocal findings were graded 0.5. These were not counted as fractures. Non-fracture deformity was not reported separately, and all positives identified by SQ were counted as fractures (and are referred to as such in this report). The SQ assessment was performed prior to development of the new approach (ABQ).

3. Algorithm-based qualitative assessment (ABQ). Reader 2 (G.J.) performed algorithm-based qualitative assessment (ABQ) of vertebral fractures. This was applied in the same study population with a time gap of 4 years after the original SQ assessment made by G.J. The ABQ analysis was performed blinded to the results of previous readings so that reader 2 was not influenced by the earlier SQ assessment. Reader 2 developed the ABQ approach by evaluating discrepancies between different methods of vertebral fracture definition, and by examining the nature of osteoporotic vertebral fracture through observation of the characteristics of incident fractures. It was considered that incident vertebral fractures are more likely to represent true fractures than those identified at baseline, because they cannot be attributed to normal variation (as this would have been identified previously). Nor are incident fractures likely to represent non-fracture deformities, because these develop over a longer time period. The observations were made from spinal radiographs in approximately 1,050 subjects. These individuals included study subjects participating in epidemiological studies and clinical trials, and patients attending the metabolic bone clinic (Northern General Hospital, Sheffield, UK). Incident vertebral fractures occurred in 86 of these subjects, and it was observed that these new fractures always without exception involved changes in the vertebral endplate.

Criteria were formulated for three types of classification: (1) normal, (2) osteoporotic fracture, and (3) non-fracture deformity, normal variant or fracture due to causes other than osteoporosis. On the basis of these criteria, a decision-making algorithm was developed for the identification of vertebral fracture (Fig. 1). Diagnosis of osteoporotic fracture in ABQ is based on the assumption that these fractures always involve fracture of the endplate within the vertebral ring (sustained when the endplate collapses under the pressure of the intervertebral disk). Thus, using this approach, evidence of endplate fracture (rather than variation in vertebral shape) is the primary indicator of osteoporotic fracture.

The ABQ approach also assumes that in every vertebral fracture, the collapse occurs primarily at the center of the endplate (this represents the weakest part of the vertebral body as it is not protected by the strong

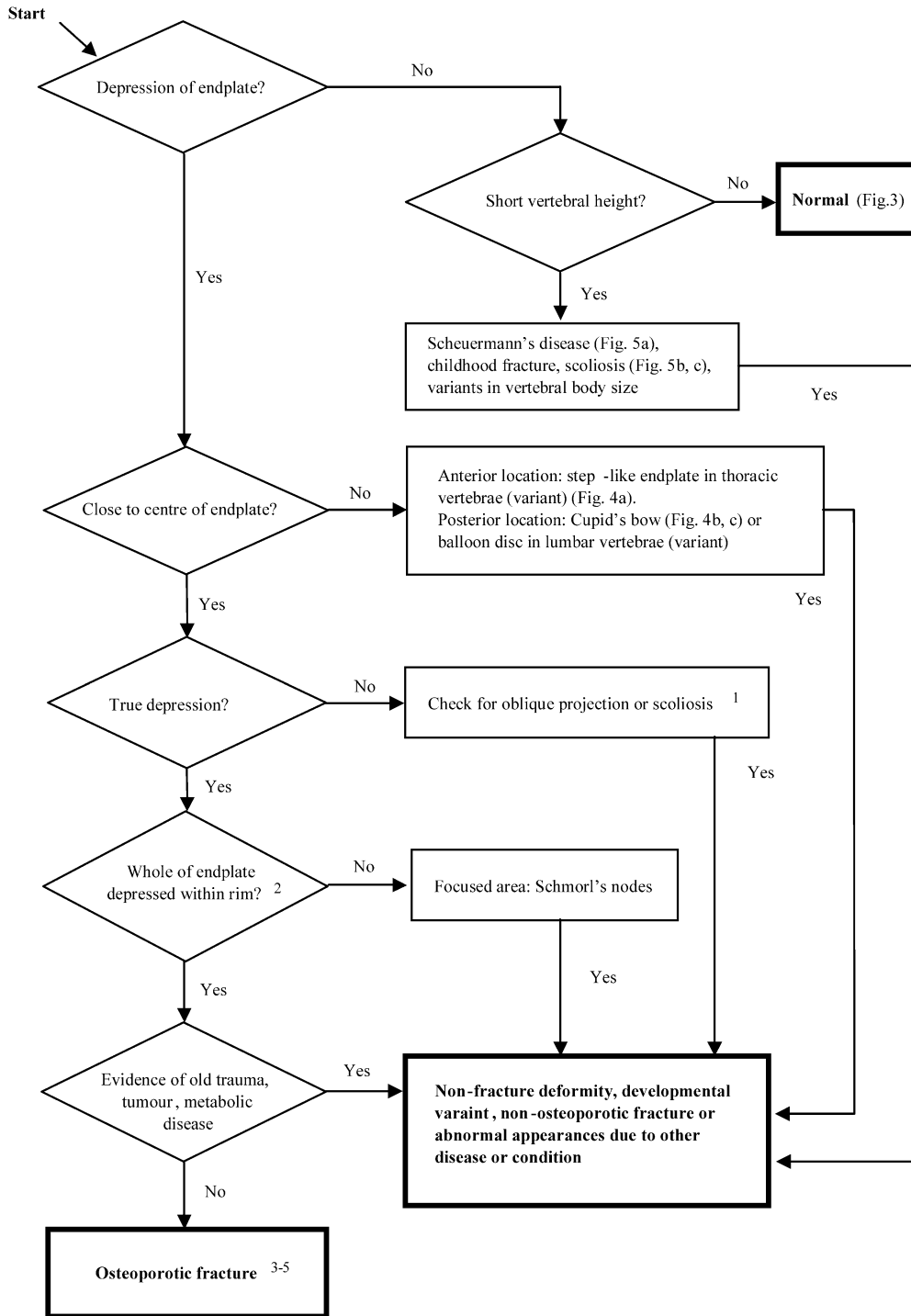
outer vertebral ring). Thus by definition, wedge and crush fractures are also concave fractures, because they involve central depression of the endplate. Concave fracture is indicated by depression of the line that represents the endplate within the vertebral ring (the intact vertebral ring may, or may not, be evident). The concavity is predominantly sited at, or close to, the center of the endplate. In wedge fracture, there is concave fracture of the endplate with fracture of the anterior cortex of the vertebral body, possible displacement of the anterior vertebral ring, and reduction in anterior height. Crush fractures involve concave fracture of the endplate plus fracture of the anterior and lateral cortices, with reduction of all three (anterior, middle, and posterior) vertebral heights. There may be buckling of the anterior surface of the vertebral body and very occasionally posterior buckling (this is not commonly seen, because the posterior portion of the vertebral body is stronger than the anterior).

For the purposes of this analysis, prevalent vertebral fractures identified by ABQ were not graded according to severity. The experienced reader may estimate the severity of fracture according to the apparent degree of vertebral height reduction, but without the aid of direct measurements, the reduction in vertebral height cannot be determined with accuracy, particularly when it is close to the cutoff for grading of fracture. Also, for prevalent fractures the original dimensions of the vertebra in question are unknown. Although this study was concerned only with the diagnosis of prevalent fractures, we have formulated the following approach for the identification of incident vertebral fractures using ABQ:

1. New fracture: change from normal appearances to the appearances of prevalent fracture according to ABQ (Fig. 1)
2. Worsening of preexisting fracture:
  - a. New fracture at the opposite endplate (in a vertebra with prevalent fracture at one endplate only)
  - b. Change in type of fracture at the same endplate (change from concave to wedge or crush, or from wedge to crush fracture)
  - c. Further reduction in vertebral height at the same endplate. In severe worsening of prevalent fracture, this will be visually obvious to the expert reader. For less obvious changes, a further reduction in height of approximately 4 mm or more is required. (Check for variation in patient positioning and magnification differences between serial radiographs when the degree of further height reduction is small or ambiguous.)

#### Comparison of methods

The numbers and distribution of fractures were compared for each of the three methods. Fractures identified as positive by all three visual methods were counted as



true. For fractures not agreed by all three methods, the radiographic appearances were evaluated in an attempt to identify the causes of discrepancy. We also compared the identification of vertebral fractures by each of the visual methods to quantitative morphometry using both the Eastell-Melton [15] and McCloskey [16] algorithms with cut points of mean  $-3$  and mean  $-4$  SD.

Because there is currently no agreed gold standard for the identification of vertebral fracture, we attempted to validate true fractures by examining the association

between vertebral fracture and BMD for each of the visual approaches. We assumed that deformities accompanied by low bone density were more likely to be true osteoporotic fractures. Expected values for BMD were calculated from measurements in the whole population (vertebrae that were fractured were excluded from analysis of lumbar spine bone density scans). Standard deviation units for BMD ( $Z$ -scores) were calculated from the expected values (to adjust for the effect of age on bone density). Age, height, weight, and BMD were



**Fig. 1** Algorithm for qualitative identification of vertebral fracture, non-fracture deformity, and normal variant. 1 In a perfect projection the endplate and vertebral ring lines are superimposed. In an oblique projection, two separate lines may represent the vertebral ring. The endplate may be superimposed on one of the vertebral ring lines, or may be seen separately as a third line. Check the posterior vertebral components for signs of rotation or scoliosis. This can give misleading appearances of endplate depression. 2 The line representing the endplate begins slightly medial to the anterior edge of the superior and inferior surfaces of the vertebral body. If the depressed line begins at the extreme anterior corner of the vertebral body, it is unlikely to be representative of fracture of the endplate. 3 Symmetry of the endplates. Osteoporotic fracture commonly involves a single endplate (superior or inferior). When both endplates are fractured, one endplate is often more severely affected, and will usually be easy to determine. If fracture is suspected, but it is unclear which endplate is depressed, check for non-fracture deformity. If fracture is suspected and the disk spaces are symmetrical, check for osteomalacia or balloon disk. 4 Type of osteoporotic fracture. Concave: fracture at the endplate alone (minor concave fracture may be missed in oblique projections); Wedge: fracture at the endplate and anterior cortex (there may be buckling at the anterior vertebral ring); Crush: fracture at the endplate and anterior and lateral-posterior cortices (there may be buckling at the anterior or posterior vertebral ring). 5 Severity of fracture: cannot be performed accurately without direct measurements. Worsening of a prevalent fracture can be determined at the time of follow-up by comparison of serial radiographs

compared in women with vertebral fractures and in those with no vertebral fractures according to each of the three methods, and for Qual and ABQ they were compared in women with non-fracture deformities or normal variants. Age, height, weight, BMD, and BMD Z-scores were compared in the women with vertebral fracture according to each of the three methods. We also compared BMD and BMD Z-scores in women with agreed vertebral fracture and in those with discrepant results for SQ versus Qual, SQ versus ABQ, and Qual versus ABQ.

**Statistical analysis**

Agreement between methods was calculated in MedCalc (MedCalc, Mariakerke, Belgium) using  $\kappa$  statistics. Mean age, weight, height, and BMD were compared in women with and without vertebral fracture using two-sample *t*-tests. Expected BMD values were calcu-

lated from the whole study population using multiple regression analysis. The BMD Z-scores were calculated as observed minus expected BMD / SD. Absolute BMD and BMD Z-scores in women with vertebral fractures identified by each of the methods were compared using one-way ANOVA, and BMD and BMD Z-scores in the women with agreed fractures and those with discrepant results were compared using two-sample *t*-tests. These analyses were performed in Statgraphics Plus for Windows version 5.0 (Statistical Graphics Corporation, Rockville, MD, USA). A statistical significance level of  $p < 0.05$  was used for all analyses.

**Results**

**Identification of vertebral fractures and deformities**

Three participants were excluded from analysis due to severe scoliosis, bringing the total number of participants studied to 372. The total numbers of vertebral fractures and women with vertebral fractures identified by each method are shown in Table 1. Almost all fractures diagnosed by ABQ were also diagnosed by SQ and Qual, whereas only a very few fractures diagnosed by ABQ were not confirmed by SQ or Qual. The prevalence of fracture (by subjects) was highest for SQ and lowest for ABQ (24% prevalence for SQ compared with 11% for Qual and 7% for ABQ). There were relatively few women identified with fracture by Qual (six women) and ABQ (one woman) that were not agreed by SQ. The total number of women with vertebral fractures agreed by all three methods was 22. There were more women identified with either non-fracture deformities or normal variants by ABQ (62 women) than by Qual (47 women).

Agreement for the identification of vertebral fracture by study participants was moderate between Qual and ABQ ( $\kappa = 0.66$ , SE = 0.07; 95% CI, 0.52 to 0.81) and was relatively poor between SQ and Qual ( $\kappa = 0.44$ , SE = 0.07; 95% CI, 0.31 to 0.57), and SQ and ABQ ( $\kappa = 0.36$ , SE = 0.07; 95% CI, 0.22 to 0.50). For analysis by vertebrae, the  $\kappa$  scores were lower (for Qual versus ABQ,  $\kappa = 0.57$ , SE = 0.06; 95% CI, 0.45 to 0.69; for SQ versus Qual,  $\kappa = 0.36$ , SE = 0.05; 95% CI, 0.27 to 0.46; and for SQ versus ABQ,  $\kappa = 0.30$ , SE = 0.06; 95% CI,

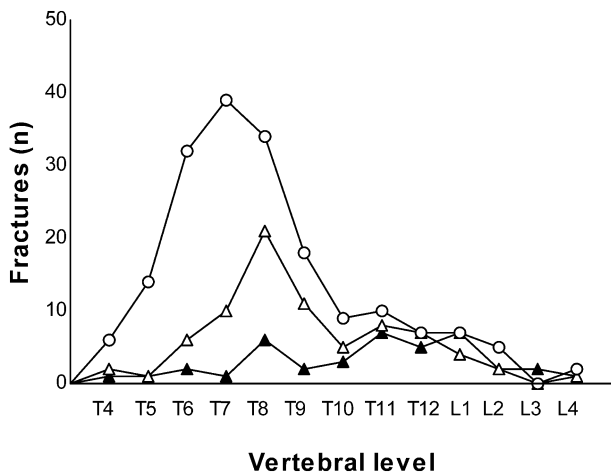
**Table 1** Visual identification of vertebral fracture using three different methods. Results show the number of women (vertebrae in parentheses) with (Yes) or without (No) vertebral fracture according to semiquantitative (SQ), qualitative (Qual), and

algorithm-based qualitative identification of vertebral fracture (ABQ) in 372 postmenopausal women. Non-fracture deformities, normal variants, and other abnormalities were diagnosed separately using Qual and ABQ but not SQ

	Fracture	SQ; subjects, <i>n</i> = 88 (vertebrae, <i>n</i> = 182)		ABQ; subjects, <i>n</i> = 25 (vertebrae, <i>n</i> = 40)	
		Yes	No	Yes	No
Qual; subjects, <i>n</i> = 39 (vertebrae, <i>n</i> = 78)	Yes	33 (49)	6 (29)	22 (34)	17 (44)
	No	55 (133)	278 (4,997)	3 (6)	330 (5,124)
ABQ; subjects, <i>n</i> = 25 (vertebrae, <i>n</i> = 40)	Yes	24 (34)	1 (6)		
	No	64 (148)	283 (5,020)		

**Table 2** Agreement between visual identification of vertebral fractures and quantitative morphometry. Results show the  $\kappa$  scores for agreement between semiquantitative (SQ), qualitative (Qual), and algorithm-based qualitative (ABQ) identification of vertebral fracture, and quantitative morphometry on a per-subject (per-vertebra) basis. The Eastell-Melton and McCloskey quantitative methods were applied with cut-points for vertebral deformity of mean - 3 SD and mean - 4 SD

QM	SQ; subjects, $n = 88$ (vertebrae, $n = 182$ )	Qual; subjects, $n = 39$ (vertebrae, $n = 78$ )	ABQ; subjects, $n = 25$ (vertebrae, $n = 40$ )
Eastell-Melton, 3SD Subjects, $n = 47$ (Vertebrae, $n = 73$ )	0.42 (0.35)	0.42 (0.39)	0.54 (0.53)
Eastell-Melton, 4SD Subjects, $n = 23$ (Vertebrae, $n = 29$ )	0.31 (0.22)	0.48 (0.39)	0.64 (0.64)
McCloskey, 3SD Subjects, $n = 73$ (Vertebrae, $n = 98$ )	0.34 (0.29)	0.36 (0.33)	0.39 (0.43)
McCloskey, 4SD Subjects, $n = 30$ (Vertebrae, $n = 37$ )	0.23 (0.20)	0.38 (0.36)	0.59 (0.54)



**Fig. 2** The distribution of prevalent vertebral fractures in 372 postmenopausal women according to vertebral level. The *open circles* represent fractures identified by semiquantitative assessment (SQ); *open triangles*, fractures identified by qualitative assessment (Qual), and *solid triangles*, fractures identified by the algorithm-based qualitative approach (ABQ)

0.19 to 0.41). When the analysis was restricted to SQ grade 2 or grade 3 fractures only (11 subjects), the  $\kappa$  score was slightly worse for SQ versus Qual ( $\kappa = 0.37$ ,  $SE = 0.11$ ; 95% CI, 0.16 to 0.59), but was better between SQ and ABQ ( $\kappa = 0.54$ ,  $SE = 0.11$ ; 95% CI, 0.31 to 0.76) than for all SQ fractures. For analysis of SQ grade 2 and grade 3 fractures by vertebrae, the  $\kappa$  scores were similar to those for analysis by subjects. The ABQ diagnosis of vertebral fracture agreed more closely with quantitative assessment of vertebral deformities (both the Eastell-Melton and the McCloskey methods) than either Qual or SQ. This was true for analysis both by subjects and by vertebrae, and for both the 3 and 4 SD cut points for deformity (Table 2).

For SQ and Qual, fractures were more common in the midthoracic region (Fig. 2). For ABQ, fractures were more common at the junction of thoracic and lumbar spine. SQ identified more fractures than both Qual and ABQ at each vertebral level between T4 and T11. Most of the discrepancies between SQ and the other approaches occurred in the midthoracic region.

### Analysis of discrepancies between methods

The causes of disagreement between methods for the identification of vertebral fracture are listed in Table 3. A large number of the 88 women with SQ fractures (55 women for Qual and 64 women for ABQ) were identified as fracture negative (normal, non-fracture deformity, normal variant, or other abnormality) by Qual and ABQ. When the numbers of women with abnormal vertebrae were compared after including women with non-fracture deformities or normal variants in the totals for Qual and ABQ, the prevalence was similar for each of the three methods (88, 80, and 86 women with abnormal vertebrae for SQ, Qual, and ABQ, respectively).

Nearly three quarters of women with SQ fractures (70%) were identified as abnormal but fracture-negative by ABQ. The largest single contributor to lack of agreement for the identification of vertebral fracture was Scheuermann's disease, followed by normal variants, and then degenerative change. Of the women with vertebral fractures identified by SQ alone, at least half of the women (52% for Qual and 63% for ABQ) had evidence of Scheuermann's disease or degenerative change with short anterior vertebral height.

### Relationship between vertebral fracture and height, weight, age, and bone density

Women with vertebral fractures identified by all three methods were older, and had lower weight and lower BMD compared with women with no vertebral fractures (Table 4). Women with vertebral fractures according to ABQ were also shorter than those with no fracture. The deficit in BMD in women with vertebral fractures was less marked for SQ than for Qual and ABQ. Mean BMD at the lumbar spine and total body was lower in women with fractures identified by Qual and ABQ than by SQ, and BMD at the trochanter was lower for ABQ than for Qual and SQ. The differences in BMD remained after adjusting for age, with lower mean BMD Z-scores at the lumbar

**Table 3** Sources of discrepancy between different methods for the identification of vertebral fracture. Results show the number of women and vertebrae (*Vert*) in a population of 372 postmenopausal women with discrepant diagnosis of vertebral fracture identified by three methods. The numbers in parentheses represent the percentage of the total number of women with vertebral frac-

ture and the total number of fractured vertebrae identified by semiquantitative (*SQ*), qualitative (*Qual*), and qualitative assessment using an algorithm-based qualitative approach (*ABQ*). Non-fracture deformities, normal variants, and other abnormalities were not diagnosed separately using the *SQ* method

	SQ: fracture (88 women)				Qual: fracture (39 women)		ABQ: fracture (25 women)	
	Qual: no fracture		ABQ: no fracture		ABQ: no fracture		Qual: no fracture	
	Women	Vert	Women	Vert	Women	Vert	Women	Vert
Scheuermann's disease	21 (24%)	46	24 (27%)	56	1 (3%)	5	0	0
Normal variant	0 (0%)	0	18 (20%)	33	11 (28%)	26	0	0
Degenerative change and short anterior height	25 (28%)	51	13 (15%)	22	3 (8%)	9	0	0
Lateral wedge (developmental) and short anterior height	0 (0%)	0	5 (6%)	10	1 (3%)	3	0	0
Old trauma	0 (0%)	0	2 (2%)	2	0 (0%)	0	0	0
Scoliosis	1 (1%)	2	0 (0%)	0	1 (3%)	1	0	0
Normal	8 (9%)	12	2 (2%)	2	0 (0%)	0	3 (12%)	4
Discrepancies (total)	55 (63%)	111	64 (73%)	125	17 (44%)	44	3 (12%)	4

spine and total body in women with vertebral fractures identified by Qual and ABQ than in those identified by SQ (Table 5).

In women with SQ fractures that were agreed by Qual and ABQ, the BMD and BMD Z-scores for the lumbar spine were significantly lower than in women with SQ fractures identified by Qual and ABQ as non-fracture deformities or normal variants ( $p < 0.01$ ), and the BMD

Z-scores for the women with agreed SQ fractures were not significantly different from zero (one-sample *t*-test) (data not shown). When the analysis was restricted to women with SQ fractures grades 2 or 3, there were no significant differences in the BMD Z-scores for the three visual methods. All of the 17 women identified with vertebral fractures by Qual, but not ABQ, had non-fracture deformities, normal variants, or evidence of old

**Table 4** Characteristics of women with vertebral fracture identified by different methods. Results show the mean values (SD) in women with and without vertebral fractures. Fractures were identified by semiquantitative (*SQ*), qualitative (*Qual*), and algorithm-based qualitative assessment of vertebral fracture (*ABQ*) in 372 postmenopausal women. Bone mineral density (*BMD*) was measured using the Lunar DPX (Lunar, Madison, WI, USA). Means were compared in women with and without vertebral fracture according to each approach using two-sample *t*-tests. Means for BMD were also compared in women with fractures according to each method

using one-way ANOVA and Fisher's least significant difference test. Values that have the same letter are not significantly different ( $p < 0.01$ ). If two values are given an (*a*) and a (*b*), respectively, those values differ significantly when all three analyses of the approaches to identification of vertebral fracture were compared. If a value has both an (*a*) and a (*b*), it is not significantly different from values with either an (*a*) or a (*b*) when all three analyses of the approaches to identification of vertebral fracture were compared

	SQ		Qual		ABQ	
	Fracture (n = 88)	No fracture (n = 284)	Fracture (n = 39)	No fracture (n = 333)	Fracture (n = 25)	No fracture (n = 347)
Age (years)	67.1** (9.1)	63.6 (8.9)	68.3** (8.9)	64.1 (9.1)	70.1*** (7.9)	64.0 (9.0)
Height (m)	1.58 (0.07)	1.59 (0.06)	1.57 (0.07)	1.59 (0.06)	1.56* (0.06)	1.59 (0.06)
Weight (kg)	63.0** (9.6)	67.2 (12.8)	61.8* (9.9)	66.6 (12.3)	60.5* (8.5)	66.6 (12.4)
BMD (g/cm <sup>2</sup> )						
Lumbar spine	(a) 1.006*** (0.195)	1.088 (0.186)	(b) 0.889*** (0.173)	1.080 (0.186)	(b) 0.884*** (0.176)	1.082 (0.185)
Femoral neck	0.794*** (0.122)	0.862 (0.134)	0.754** (0.100)	0.852 (0.134)	0.746*** (0.099)	0.863 (0.134)
Ward's triangle	0.633*** (0.146)	0.750 (0.168)	0.615*** (0.113)	0.737 (0.167)	0.605*** (0.113)	0.739 (0.167)
Trochanter	(a) 0.679*** (0.116)	0.764 (0.134)	(ab) 0.633*** (0.101)	0.751 (0.134)	(b) 0.625*** (0.101)	0.752 (0.133)
Total body	(a) 1.027*** (0.111)	1.083 (0.103)	(b) 0.962*** (0.089)	1.077 (0.105)	(b) 0.957*** (0.085)	1.078 (0.104)

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; (ANOVA) for comparison of values for comparison of means in women with and without vertebral fractures according to each method (2 sample *t*-tests)

**Table 5** Comparison of bone mineral density Z-scores in women with vertebral fractures identified by different methods. Results show the mean bone mineral density (BMD) Z-scores (age-adjusted SD units for BMD) in women with vertebral fractures identified by semiquantitative (SQ), qualitative (Qual), and algorithm-based qualitative assessment of vertebral fracture (ABQ), in 372 postmenopausal women. Bone mineral density was measured in g/cm<sup>2</sup> using the Lunar DPX (Lunar, Madison, WI, USA). Means were compared using one-way ANOVA and Fishers least significant difference test. Values that have the same letter are not significantly different ( $p < 0.01$ ). If two values are given an (a) and a (b), respectively, those values differ significantly when all three analyses of the approaches to identification of vertebral fracture were compared

BMD Z-score	SQ	Qual	ABQ
Lumbar spine	-0.20 (a)	-0.71 (b)	-0.75 (b)
Femoral neck	-0.22 (a)	-0.45 (a)	-0.42 (a)
Ward's triangle	-0.25 (a)	-0.47 (a)	-0.45 (a)
Trochanter	-0.35 (a)	-0.62 (a)	-0.63 (a)
Total body	-0.21 (a)	-0.78 (b)	-0.76 (b)

trauma according to ABQ. Although there was a trend toward higher absolute BMD and BMD Z-scores (at all measurement sites) in these 17 women compared to those with fractures agreed by both Qual and ABQ, the differences were not statistically significant at  $p < 0.05$  (data not shown). For both Qual and ABQ, mean age, height, weight, and BMD in women with non-fracture deformity were similar to values in the whole study population (data not shown).

## Discussion

In this study, the number of women with vertebral fractures according to SQ was twice the number identified by Qual and more than three times the number identified by ABQ. The prevalence of vertebral fracture was lowest for ABQ: this was not unexpected and we attribute this to the application of stringent criteria in ABQ to rule out deformation due to causes other than osteoporotic fracture. Agreement between the three visual methods was moderate (Qual versus ABQ) to poor (Qual versus SQ and ABQ versus SQ), and of the three visual methods, ABQ had the best consensus with quantitative morphometry. The main source of disagreement between visual methods was the large number of fractures identified by SQ, but not

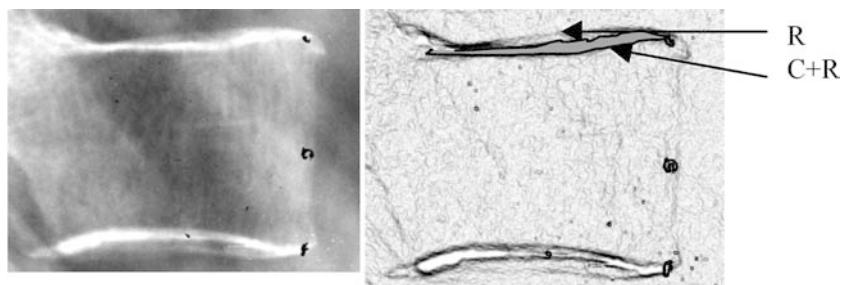
Qual or ABQ: many of these were mild thoracic wedge fractures identified as non-fracture deformities or normal variants by the other two methods. Bone density was lower in the women with fractures agreed by all three methods, than in women with discrepant SQ fractures.

Differences in the way various methods of fracture definition are applied will inevitably lead to a certain amount of disagreement between methods, but the disproportionately large number of discrepant SQ fractures identified in our study raises questions about the true nature of these fractures. While in some comparison studies, the number of vertebral fractures identified by SQ is similar to or slightly higher than the number of QM deformities [17, 18, 19], other researchers have identified a much larger proportion of fractures using SQ [20, 21, 22]. Black et al. for example, found that SQ identified three times the number of prevalent mild vertebral fractures than were identified using four different quantitative approaches, and the associations between these fractures and clinical outcomes of vertebral fracture were closer to those for normal subjects than for those with moderate or severe fractures [21].

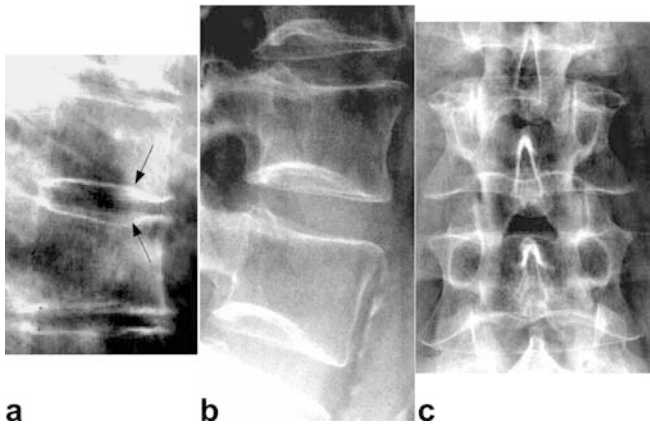
We applied SQ in accordance with the original methodology described by Genant et al. [10]. This does not include any kind of systematic approach for the identification of non-fracture deformities or normal variants. Our analysis of discrepancies between methods and our bone density data suggest that fractures may be overdiagnosed by SQ because as with QM, the method fails to differentiate adequately between true fracture and non-fracture deformity. In EVOS, for example, morphometric vertebral deformities that were fracture-negative according to radiology reports, were not associated with low bone density. This raises the likelihood that these were non-fracture deformities identified by QM on the basis of "short" vertebral height [23]. Such variation in vertebral height may be due not only to fracture, but also to developmental abnormalities with or without degenerative changes (characterized for example by small or wedged vertebral bodies and deep or step-like endplates, or by wedged and elongated vertebral bodies, irregular sclerotic endplates and Schmorl's nodes due to adolescent Scheuermann's and/or degenerative disk disease) (Figs. 3, 4, and 5).

The discrepant SQ fractures in our study were predominantly mild wedge fractures in the midthoracic

**Fig. 3** Appearance of vertebral endplates in a normal vertebra. R represents the vertebral ring line, C + R represent the central endplate within the vertebral ring overlapping the vertebral ring line

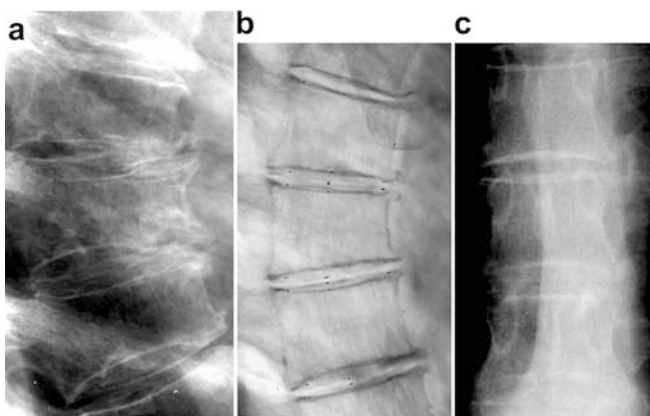






**Fig. 4** **a** Developmental deformity. The inferior endplate of vertebra T7 and the superior endplate of T8 are deeper anteriorly (arrows). These endplates appear symmetrical about the intervertebral disk. **b** The endplate is much deeper (arrow) at the posterior end of the inferior endplate of L3. **c** The AP view shows that the change is caused by a 'Cupid's bow' developmental deformity

region. The clinical significance of mild vertebral fractures or deformities remains controversial [24, 25], particularly for wedge fractures. Evidence from one study, for example, suggests that osteoarthritis is associated with short anterior height only, whereas in patients with osteoporotic fracture, both middle and anterior heights are affected [26], and elsewhere, vertebral endplate deformities have been shown to be more strongly associated with low bone density than wedge deformities [27]. The differentiation between wedge deformity and fracture is not easy, due to our limited knowledge about the causes of gradual wedging and the appearance of very early fracture [28]. The wedge fractures that were



**Fig. 5** **a** Non-fracture deformity: there is short height of vertebrae T7 and T9, but no fracture of the endplate within the vertebral ring. The endplate and the vertebral rings are irregular, and indicate Scheuermann's disease. **b** There is short anterior height in vertebra T8 but no endplate fracture. The vertebral ring lines in both the superior and inferior endplates are straight and are converging anteriorly. **c** In the AP view, vertebra T8 has a slight lateral wedge on the right side (developmental)

identified by SQ but not ABQ in our study were questionable, because none of them exhibited central depression of the endplate.

In the absence of a gold standard, we cannot confirm the true nature of discrepancies between methods, but we can confidently validate the diagnosis in the 22 women with fractures agreed by all three methods. Most of the women with ABQ fractures (22 out of 25 women) were among the women identified by all three methods, and the prevalence of vertebral fracture according to ABQ (7%) was similar to that obtained by visual reading in the European Vertebral Osteoporosis Study (5.7%) [23]. Among the 17 women with vertebral fractures identified by Qual but not ABQ, there was a trend toward higher BMD, and all of the discrepant Qual fractures were identified as non-fracture deformities or normal variants by ABQ.

The lack of a gold standard for the identification of vertebral fractures and the general lack of understanding of the development of vertebral fracture and non-fracture deformity are the main limitations of this study. This was a preliminary evaluation of a modified approach to the qualitative identification of vertebral fracture, and although the ABQ method shows promise, we do not as yet have sufficient evidence to advocate ABQ as a replacement for other visual radiological methods. However we are optimistic that because the approach is standardized, interobserver agreement between readers may be improved, and that this report will stimulate further debate and closer examination of the influence of non-fracture deformities and normal variants on the identification of vertebral fractures.

**Acknowledgements** This study was supported by a program grant from the Arthritis Research Campaign. Dr Diana Greenfield was responsible for recruiting subjects, and we thank the staff of the Osteoporosis Centre, Northern General Hospital, Sheffield, UK, in particular the radiographers, for acquisition of BMD scans. Thanks also to Dr Nicola Peel and Dr Jackie Clowes for their assistance in the development of ABQ.

## References

- Hedlund LR, Gallagher JC (1988) Vertebral morphometry in diagnosis of spinal fractures. *Bone Miner* 5:59-67
- Minne HW, Leidig G, Wuster C, Siromachkostov L, Baldauf G, Sauer P, Lojen M, Ziegler R (1988) A newly developed spine deformity index (SDI) to quantitate vertebral crush fractures in patients with osteoporosis. *Bone Miner* 3:335-349
- Smith-Bindman R, Steiger S, Cummings SR, Genant HK (1991) The index of radiographic area (IRA): a new approach to estimating the severity of vertebral deformity. *Bone Miner* 15:137-150
- Black DM, Cummings SR, Stone K, Huden K, Palermo L, Steiger P (1991) A new approach to defining vertebral dimensions. *JBMR* 8:883-891
- Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ III (1991) Classification of vertebral fractures. *J Bone Miner Res* 6:207-215
- Davies KM, Recker RR, Heaney RP (1993) Revisable criteria for vertebral deformity. *Osteoporos Int* 3:265-270

7. McCloskey EV, Spector TD, Eyres KS, Fern ED, O'Rourke N, Vasikaran S, Kanis JA (1993) The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* 3:138–147
8. Melton LJ, Lane AW, Cooper C, Eastell R, O'Fallon WM, Riggs BL (1993) Prevalence and incidence of vertebral deformities. *Osteoporos Int* 3:113–119
9. Ross PD, Yhee YK, He Y-F, Davis JW, Kamimoto C, Epstein RS, Wasnich RD (1993) A new method for vertebral fracture diagnosis. *JBMR* 2:167–174
10. Genant K, Wu CY, Kuijk CV, Nevitt MC (1993) Vertebral fracture assessment using a semi-quantitative technique. *J Bone Miner Res* 8:1137–1148
11. Hurxthal LM, Vose GP, Detter WE (1969) Densitometric and visual observations of spinal radiographs. *Geriatrics* 24:93–106
12. Jensen FG, McNair P, Boesen J et al (1984) Validity in diagnosing osteoporosis. *Eur J Radiol* 4:1–3
13. O'Neill TW, Felsenberg D, Varlow J, Cooper C, Kanis JA, Silman AJ (1996) The prevalence of vertebral deformity in European men and women: the European Vertebral Osteoporosis Study. *J Bone Miner Res* 11:1010–8
14. Peel NFA, Barrington NA, Blumsohn A, Colwell A, Hannon R, Eastell R (1995) Bone mineral density and bone turnover in spinal osteoarthritis. *Ann Rheum Dis* 54:867–871
15. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ III (1991) Classification of vertebral fractures. *J Bone Miner Res* 6:207–215
16. McCloskey EV, Spector TD, Eyres KS, Fern ED, O'Rourke N, Vasikaran S, Kanis JA (1993) The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* 3:138–147
17. Grados F, Roux C, de Vernejoul MC, Utard G, Sebert JL, Fardellone P (2001) Comparison of four morphometric definitions and a semiquantitative consensus reading for assessing prevalent vertebral fractures. *Osteoporos Int* 12:716–722
18. Szulc P, Munoz F, Sornay-Rendu E, Paris E, Souhami E, Zanchetta J, Bagur A, Van der Mooren MJ, Young S, Delmas PD (2000) Comparison of morphometric assessment of prevalent vertebral deformities in women using different reference data. *Bone* 27:841–846
19. Pak CYC, Ho A, Poindexter J, Peterson R, Sakhaee K (1996) Quantitation of incident spinal fractures: comparison of visual detection with quantitative morphometry. *Bone* 18:349–353
20. Genant HK, Jergas M, Palermo L, Nevitt M, Valentin RS, Black D and Cummings SR (1996) Comparison of semiquantitative visual and quantitative morphometric assessment of prevalent and incident vertebral fractures in osteoporosis. *J Bone Miner Res* 11:984–996
21. Black DM, Palermo L, Nevitt MC, Genant HK, Epstein R, San Valentin R, Cummings SR (1995) Comparison of methods for defining prevalent vertebral deformities: the Study of Osteoporotic Fractures. *J Bone Miner Res* 16:890–902
22. Wu CY, Li J, Jergas M, Genant HK (1995) Comparison of semiquantitative techniques for the assessment of prevalent and incident vertebral fractures. *Osteoporos Int* 5:354–370
23. Leidig-Bruckner G, Limberg B, Felsenberg D, Bruckner T, Holder S, Kather A, Mikasch J, Wuster C, Ziegler R, Scheidt-Nave C (2000) Sex difference in the validity of vertebral deformities as an index of prevalent vertebral osteoporotic fractures: a population survey of older men and women. *Osteoporos Int* 11:102–119
24. Spector TD, McCloskey EV, Doyle DV, Kanis JA (1993) Prevalence of vertebral fracture in women and the relationship with bone density and symptoms: the Chingford Study. *Osteoporos Int* 8:817–822
25. Kleerekoper M (1992) Vertebral fracture or vertebral deformity? *Calcif Tissue Int* 50:5–6
26. Abdul-Hamid Osman A, Bassiouni H, Koutri R, Nus J, Geusens P, Dequeker J (1994) Aging of the thoracic spine: distinction between wedging in osteoarthritis and fracture in osteoporosis—a cross sectional and longitudinal study. *Bone* 15:437–442
27. Jones G, White C, Nguyen T, Sambrook PN, Kelly PJ, Eisman JA (1996) Prevalent vertebral deformities: relationship to bone mineral density and spinal osteophytosis in elderly men and women. *Osteoporos Int* 6:233–239
28. Lyritis GP, Mayasis B, Tsakalacos N, Lambropoulos A, Gazi S, Karachalios T, Tsekoura M, Yiatzides A (1989) The natural history of the osteoporotic vertebral fracture. *Clin Rheumatol* 8[Suppl 2]:66–69