

Review of radiological scoring methods of osteoporotic vertebral fractures for clinical and research settings

Ling Oei · Fernando Rivadeneira · Felisia Ly ·
Stephan J. Breda · M. Carola Zillikens ·
Albert Hofman · André G. Uitterlinden ·
Gabriel P. Krestin · Edwin H. G. Oei

Received: 3 April 2012 / Revised: 16 July 2012 / Accepted: 17 July 2012 / Published online: 15 August 2012
© European Society of Radiology 2012

Abstract

Background Osteoporosis is the most common metabolic bone disease; vertebral fractures are the most common osteoporotic fractures.

Methods Several radiological scoring methods using different criteria for osteoporotic vertebral fractures exist. Quantitative morphometry (QM) uses ratios derived from direct vertebral body height measurements to define fractures. Semi-quantitative (SQ) visual grading is performed according to

height and area reduction. The algorithm-based qualitative (ABQ) method introduced a scheme to systematically rule out non-fracture deformities and diagnoses osteoporotic vertebral fractures based on endplate depression. The concordance across methods is currently a matter of debate.

Results This article reviews the most commonly applied standardised radiographic scoring methods for osteoporotic vertebral fractures, attaining an impartial perspective of benefits and limitations. It provides image examples and

L. Oei · F. Rivadeneira · F. Ly · S. J. Breda · M. C. Zillikens ·
A. G. Uitterlinden
Department of Internal Medicine, Erasmus Medical Center,
Rotterdam, The Netherlands

L. Oei
e-mail: h.l.d.w.oei@erasmusmc.nl

F. Rivadeneira
e-mail: f.rivadeneira@erasmusmc.nl

F. Ly
e-mail: felisialy@gmail.com

S. J. Breda
e-mail: stephanbreda@gmail.com

M. C. Zillikens
e-mail: m.c.zillikens@erasmusmc.nl

A. G. Uitterlinden
e-mail: a.g.uitterlinden@erasmusmc.nl

L. Oei · F. Rivadeneira · F. Ly · S. J. Breda · A. Hofman ·
A. G. Uitterlinden
Department of Epidemiology, Erasmus Medical Center,
P.O. Box 2040 Ee21-75, 3000 CA Rotterdam, The Netherlands

A. Hofman
e-mail: a.hofman@erasmusmc.nl

L. Oei · F. Rivadeneira · M. C. Zillikens · A. Hofman ·
A. G. Uitterlinden
Netherlands Genomics Initiative (NGI)-sponsored Netherlands
Consortium for Healthy Aging (NCHA),
Rotterdam, The Netherlands

G. P. Krestin · E. H. G. Oei (✉)
Department of Radiology, Erasmus Medical Center,
's Gravendijkwal 230,
3015 CE Rotterdam, The Netherlands
e-mail: e.oei@erasmusmc.nl

G. P. Krestin
e-mail: e.oei@erasmusmc.nl

L. Oei · F. Ly · S. J. Breda
Departments of Internal Medicine and Epidemiology,
Erasmus Medical Center,
P.O. Box 2040 Ee21-83, 3000 CA Rotterdam, The Netherlands

F. Rivadeneira
Departments of Internal Medicine and Epidemiology,
Erasmus Medical Center,
P.O. Box 2040 Ee5-79, 3000 CA Rotterdam, The Netherlands

M. C. Zillikens
Department of Internal Medicine, Erasmus Medical Center,
's Gravendijkwal 230,
3015 CE Rotterdam, The Netherlands

A. G. Uitterlinden
Departments of Internal Medicine and Epidemiology,
Erasmus Medical Center,
P.O. Box 2040 Ee5-75B, 3000 CA Rotterdam, The Netherlands

discusses aspects that facilitate large-scale application, such as automated image analysis software and different imaging investigations. It also reviews the implications of different fracture definitions for scientific research and clinical practice. *Conclusion* Accurate standardised scoring methods for assessing osteoporotic vertebral fractures are crucial, considering that differences in definition will have implications for patient care and scientific research. Evaluation of the feasibility and concordance among methods will allow establishing their benefits and limitations, and most importantly, optimise their effectiveness for widespread application.

Key Points

- Several scoring methods using different criteria for assessing osteoporotic vertebral fractures exist.
- Standardised osteoporotic vertebral fracture assessment should be applicable to different radiological investigations.
- Accurate assessment of osteoporotic vertebral fractures is essential for proper patient management.
- Optimising feasibility of scoring methods enables widespread use in scientific research.
- Assessment of concordance between methods is important for application in patient care.

Keywords Osteoporosis · Vertebral Fracture · Radiography · Diagnosis · Epidemiologic Studies

Abbreviations

BMD	Bone mineral density
QM	Quantitative morphometry
SQ	Semi-quantitative method
SFI	Spinal fracture index
ABQ	Algorithm-based qualitative method
DXA	Dual-emission X-ray absorptiometry
SD	Standard deviation
FRAX	Fracture risk assessment tool
VFA	Vertebral fracture assessment
CT	Computed tomography
MRI	Magnetic resonance imaging

Introduction

Osteoporosis is the most common metabolic bone disease, and vertebral fractures are the most common type of osteoporotic fractures [1]. These fractures are associated with significant morbidity [2–7], mortality [8, 9], and high health-care costs. Given the ageing of populations, osteoporotic vertebral fractures are likely to become an increasingly important health issue. The costs of osteoporotic vertebral fractures were estimated to be €1.5 billion in Europe in 2010 [10] and US\$ 1.1 billion in the US in 2005, and they are expected to have increased by more than 50% by 2025 [11].

The aetiology of osteoporotic vertebral fractures is believed to be multi-factorial, influenced by genetic and environmental factors [12, 13]. Osteoporosis is a disease characterised by low bone mass and micro-architectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk [14]. Bone mineral density (BMD) and age are strongly predictive for most osteoporotic fractures, whereas other risk factors may vary according to fracture site [15]. Compared with non-vertebral fractures, which usually occur after a fall, it has been suggested that only ~10–15% of osteoporotic vertebral fractures are preceded by a fall, with many resulting from low-grade trauma apparently derived from insignificant everyday activities [16].

Importantly, vertebral fractures are strong predictors of future osteoporotic fractures, including both non-vertebral and new vertebral fractures [17–20]. Vertebral fractures can be relatively asymptomatic in some cases, still, asymptomatic vertebral fractures remain strong predictors of subsequent risk of fractures and fracture-associated mortality [21]. The reason why radiological detection may be even more valuable. In clinical practice, therefore, prevalent osteoporotic vertebral fractures are considered as a strong indication for anti-osteoporotic treatment [22]. Yet, previous studies have shown that only one third of the patients with vertebral fractures come to clinical attention [23] and that vertebral fractures are commonly underreported in radiological practice [24–27]. The latter implies that applying standardised assessment methods of osteoporotic vertebral fractures might be beneficial to decrease reader subjectivity.

Currently, there is no gold standard for osteoporotic vertebral fracture diagnosis [28]. Several radiological scoring methods for osteoporotic vertebral fractures exist, each using different criteria for diagnosing and grading the fracture. Such grading definitions are currently under debate. This article will review different scoring methods for diagnosing osteoporotic vertebral fractures by discussing the benefits and limitations of the most commonly applied radiographic scoring methods. We will also discuss the role of alternative imaging techniques for assessing these fractures. In addition, this review will illustrate how prevalence of osteoporotic vertebral fractures is influenced by different scoring methods. Finally, application of scoring methods in research and patient care will be discussed.

Vertebral fracture assessment by radiography

Radiography is the standard imaging modality used for initially assessing vertebral fractures. Usually separate anteroposterior and lateral projections of the thoracic and lumbar spine are acquired, sometimes supplemented by additional views focused at the thoracolumbar junction. However, in

the scientific research setting, occasionally only lateral radiographs are obtained. Radiographic capture is rapid, image quality is mostly high, and the radiation dose is relatively low, especially compared to CT. It is important that the spine is positioned parallel to the table to enable good assessment of vertebral endplates. Imperfect centering and collimation of the x-ray beam may, however, cause oblique projection and incorrect exposure, resulting in poor image quality [29]. Also, because the x-ray beam is conical, oblique projection is worst at the film areas furthest from the centre. This distortion may hinder correct appraisal of vertebral body shape and can in some situations wrongly suggest a biconcave shape [30]. Because of superimposition of the overlying shoulder girdle, the upper thoracic spinal region cannot be clearly visualised in many cases. The iliac wings of the pelvis can also exert a similar hindrance effect on images of the lower lumbar spine. In addition, superposition of the ribs and pulmonary vasculature on the thoracic vertebrae may occasionally confound the vertebral body margins on the image.

Scoring methods

Measurements of vertebral shape

The first published standardised assessment methods use quantitative morphometry (QM), which entails direct measurement of vertebral body shape. With six-point morphometry, points are placed in the superior and inferior endplates at the anterior, middle and posterior aspects of the vertebral body. For example, two of the more recent and commonly applied QM scoring methods are those described by Eastell-Melton [31] and McCloskey-Kanis [32]. In these methods anterior, central, and posterior vertebral body heights are first measured on a lateral radiograph, and ratios between these heights are calculated. These ratios are then used to classify vertebral fractures, using cutoff values based on standard deviation reductions from normal-population means derived from epidemiological studies (Fig 1). Depending on which of the three heights are diminished, the Eastell-Melton [31] method distinguishes three types of fractures (i.e. wedge, biconcavity, or compression), and the McCloskey-Kanis method further classifies the wedge type into anterior and posterior [32]. Using ratios instead of absolute heights is preferable, as anatomical structures farther away from the film may be falsely magnified, depending on the distance of the x-ray tube from the subject. Also, vertebral height is partly associated with a person's body height [31]. It is considered essential to appraise these ratios relative to population reference data, as it has been shown that the derived vertebral height ratios are normally (Gaussian) distributed [32]. In addition, several of these methods relate the

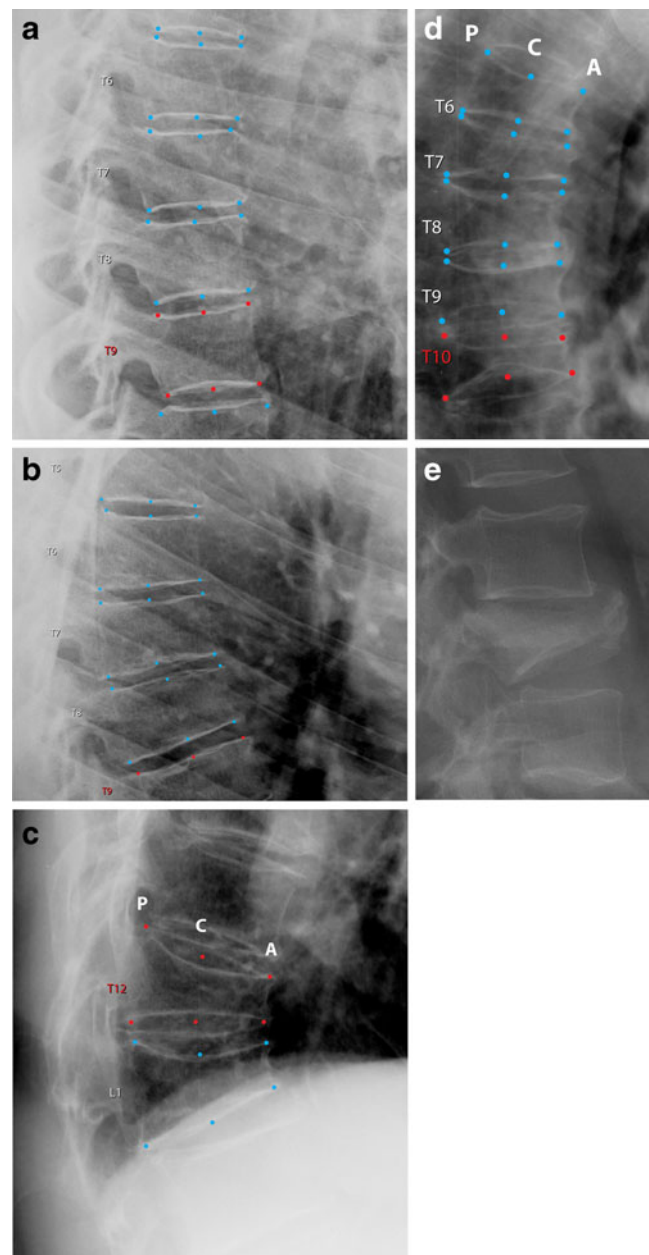


Fig. 1 Six-point quantitative morphometry (QM) and semi-quantitative (SQ) method. P, posterior; C, central; A, anterior. **a** Normal thoracic vertebrae. **b** Mild wedge deformity of T8 and **c** severe wedge deformities of T12 and L1. **d** Mild wedge deformity of T6, moderate wedge deformity of T7 and moderate biconcave deformities of T8, T9 and T10. **e** Crush deformity of L3 in an individual with confirmed history of spinal trauma; severe vertebral body fracture is seen with slight bulging of the posterior vertebral body margin. This fracture morphology is usually traumatic

values to adjacent vertebra within the same individual, as each vertebra has a different size [33]. Although QM measurements appear more objective and reproducible than visual methods, they are more laborious and time-consuming to acquire. This is an important consideration for large-scale epidemiological research as well as for its implementation in clinical practice.

Semi-quantitative (SQ) method

Currently, the most widely used standardised grading method is the visual SQ method (i.e. according to Genant [34]). It is commonly applied as a surrogate gold standard in research [35]. Vertebral fractures are SQ graded by trained readers, who estimate the percentage of height and/or area reduction subjectively, without direct measurement. Vertebral deformities are graded according to shape and severity (Fig 1). The deformity's shape is classified on the basis of anterior height loss (i.e. wedge), middle height loss (i.e. biconcave), or posterior and anterior height loss (i.e. crush). Severity of vertebral deformities is graded according to the extent of height and area loss, as mildly deformed, moderately deformed, and severely deformed (Table 1). Next, a spinal fracture index (SFI) can be calculated by summing the individual vertebral body grades. Genant et al. [34] have noted that height and area loss determined by morphometry alone fails to capture several other important characteristics of vertebral fracture, including endplate deformity, buckling of cortices, lack of parallelism of endplates and loss of vertical continuity of vertebral morphology.

Algorithm-based qualitative (ABQ) method

The more recent ABQ method by Jiang et al. diagnoses osteoporotic vertebral fractures on the basis of endplate depression, regardless of vertebral height reduction (Fig 2) [36]. The key assumption is that the endplate is always deformed in vertebral fractures, and therefore endplate depression has perfect specificity for vertebral fracture. The fracture occurs primarily at the centre of the endplate, and thus it follows that the endplate is centrally depressed in all types of vertebral fracture (i.e. concave, wedge, and crush). In addition, vertebral height may appear to be decreased as a result of oblique image projection, certain diseases, and anatomical variants that can mimic vertebral fractures. To deal with this misclassification, ABQ uses a flowchart to systematically rule out non-fracture deformities by examining

Table 1 Semi-quantitative grading of severity of vertebral fractures according to Genant [34]

Fracture severity	Grade	Reduction of:	
		Height*	Area
Normal	0		
Uncertain or borderline	0.5		
Mild	1	20–25%	10–20%
Moderate	2	25–40%	20–40%
Severe	3	≥40%	≥40%

*Anterior, middle, and/or posterior height

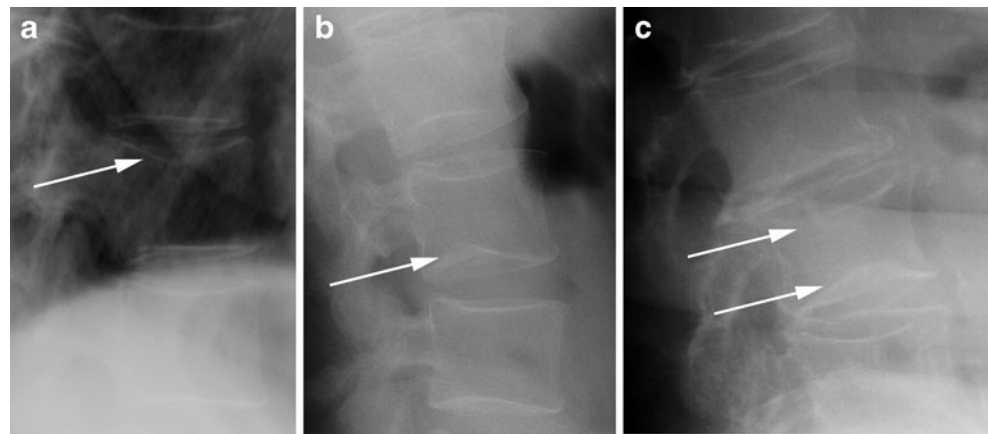
certain radiological features. A skilled ABQ reader is needed to differentiate accurately between vertebral fractures and non-fracture deformities. If images are of poor quality, vertebral fractures with subtle endplate changes can easily be missed.

Non-radiographic imaging techniques

Vertebral fractures can also be detected and graded on radiological imaging investigations other than conventional radiography. Although developed for radiography, which remains the most commonly used technique for vertebral fracture assessment, the scoring methods described above can also be applied to other radiological techniques.

Dual-energy X-ray absorptiometry (DXA) at the lumbar spine and hip to measure BMD is a routine investigation in osteoporosis, because BMD constitutes one of the strongest predictors of future fracture [37, 38]. Several studies have shown that the risk of incident vertebral fractures doubles for each SD reduction of lumbar spine BMD [37, 39]. Note, however, that many fractures occur when BMD is in the osteopenic or normal range of values [37]. In addition to artefacts (i.e. osteophytes, calcifications), BMD measurements of the lumbar spine may be falsely elevated in the presence of vertebral fractures because impacted fracture or fracture healing result in higher areal BMD [40]. The World Health Organisation's FRAX[®] tool can be used to calculate the 10-year fracture risk for individual patients, using validated risk factors (with or without femoral neck DXA BMD) [41]. The clinical risk factors used in the calculation include age, gender, height, weight, previous low trauma fracture (including vertebral fractures), parental hip fracture, oral glucocorticoid therapy, rheumatoid arthritis, current smoking, alcohol consumption of more than three units per day, and secondary causes of osteoporosis. In recent years, the use of densitometers has extended beyond BMD assessment to identify vertebral fractures from DXA images. The so-called lateral densitometric vertebral fracture assessment (VFA) is gaining popularity because of the considerable improvement in image resolution, and is currently offering complementary and independent information about fracture risk (Fig 3a) [42]. The implementation of fan-beam technology in the DXA devices has allowed capturing the whole spine in one image, with virtually no divergent beam issues due to parallax effect. Also, VFA has a low radiation dose, making it very suitable for screening in the clinical setting. For those more recently introduced DXA devices with a rotating C-arm, the lateral examination can even be done without moving the patient from the supine position used for the BMD measurements. In addition, the rotating C-arm may enable three-dimensional DXA scans, allowing the direct measurement of geometric parameters of the vertebrae

Fig. 2 Algorithm-based qualitative (ABQ) method. **a** Superior endplate depression of T11. **b** Inferior endplate depression of L3. **c** Superior and inferior endplate depression of L3



[43]. It has been demonstrated that image quality can differ greatly between types of densitometers [44]. Still, radiographs have superior spatial resolution, which facilitates identification of more subtle abnormalities [45].

Unlike two-dimensional radiography, computed tomography (CT) and magnetic resonance imaging (MRI) offer three-dimensional visualisation of the vertebra (Fig 3b). In addition, CT and MRI can differentiate between old and recent vertebral fractures by assessing the integrity and shape of the cortical margins (Fig 3c). MRI does not use ionising radiation and can demonstrate bone marrow oedema, which distinguishes recent from old fractures (Fig 3d). The images produced by CT have a much higher spatial resolution than those of MRI and DXA. It has been shown that sagittal reformations need to be used to demonstrate vertebral fractures on CT [21]. Despite the introduction of several dose reduction techniques, the ionising radiation exposure of CT is still substantial, which is a major disadvantage of the imaging technique, especially in the research setting [46]. CT scout images may also be used for assessing vertebral fractures [47]. Novel quantitative and high-resolution CT techniques are being developed to enable separate analysis of trabecular and cortical bone compartments [48, 49]. High-resolution

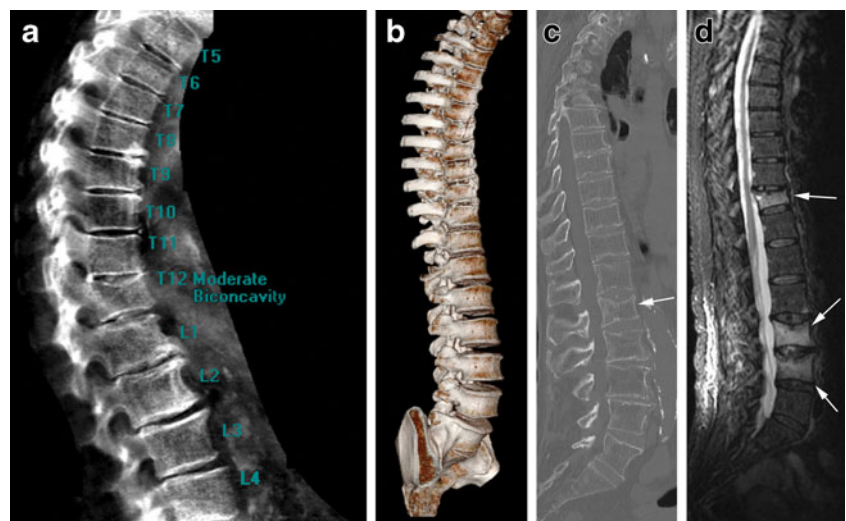
MRI can be used to assess bone trabeculation in the extremities, but this application of MRI at the spine is more challenging [50]. Drawbacks of MRI are, however, the long imaging time and high costs. Hence, MRI is usually used for other conditions that specifically require MRI, such as spinal cord compression and paraspinal soft tissue abnormalities [51, 52]. If a malignant aetiology for vertebral fracture is suspected, then MRI or CT have advantages over conventional radiography [53, 54].

Image analysis software

Automated image analysis software packages (e.g. SpineAnalyzer®, Optasia Medical Ltd, Cheadle, UK [55]) have been developed to facilitate efficient and standardised vertebral fracture scoring of large data sets [56, 57]. The software can handle lateral spine radiographs, VFA or CT scout films. So far, software packages have been dedicated to recording QM and SQ.

Instead of having to manually define vertebral contours and height, users only need to place one point in the centre of each vertebra to define the vertebral level. Next, vertebral contours

Fig. 3 Non-radiographic imaging modalities. **a** Lateral VFA shows a biconcave deformity of T12. **b** Three-dimensional visualisation of the thoracolumbar spine with CT. **c** Midline sagittal CT reformation shows an osteoporotic vertebral fracture of L1 (arrow), in addition to degenerative changes and endplate irregularities at multiple levels. **d** MRI: Sagittal short tau inversion recovery (STIR) sequence shows endplate deformity, height reduction, and bone marrow oedema at the T11, L3, and L4 levels (arrows), indicating recent osteoporotic vertebral fractures



are identified by the software using automated segmentation techniques, and vertebral height is measured (Fig 4a). A table with percentage height loss and presence or absence of deformity per vertebral level based on QM and SQ is generated (Fig 4b), after which the data can be exported to a database. The data that can be saved include the exact coordinates of the endplates anteriorly, centrally, and posteriorly. This information can be valuable in the research setting, where analysis of crude vertebral heights could be meaningful to explore optimisation of current vertebral fracture definitions. Relatively inexperienced users are deemed to be capable of using the software after a brief training [58]. These software packages can be further improved by incorporating population reference data for QM and by reducing the need for manual adjustment of vertebral contour definition, a procedure that is still required routinely. Automated VFA packages are nowadays integrated in DXA equipment.

Differential diagnosis

There are a number of differential diagnoses that have to be considered in individuals with vertebral deformities [59]. In the 1960s, Hurxthal described several criteria for vertebral measurements [30]. Basically, all artefacts that can interfere with vertebral height measurement should be considered by the reader. Hook-shaped protuberances at the posterosuperior (called uncinat process by some) and posteroinferior borders of the vertebrae, any Schmorl’s nodes, and osteophytes should be excluded from vertebral height measurement. Six-point morphometry alone is unable to distinguish fractures and vertebral deformities due to other causes. In the description of the SQ method [34], several conditions that can mimic vertebral fracture such as scoliosis and vertebral body remodelling due to degenerative disc disease are listed. Moreover, the ABQ method introduced a very comprehensive decision-

making algorithm, which provides a guideline for systematically assessing various non-fracture deformities [35].

Normal anatomical variation in the shape of individual vertebrae and of the spinal column as a whole should be taken into account. Viewed laterally, the spine has a natural curvature. Vertebrae in the mid-thoracic region are more wedge-shaped, causing a mild kyphosis. Lumbar vertebrae tend to be biconcave rather than wedge-shaped, and this gives rise to a normal lordotic curve, because of the relatively shorter posterior height. In addition, some adults have vertebrae that have longstanding short anterior height in developmentally small thoracic vertebrae [35]. Therefore, the normal spine shape must be known if SQ and ABQ readers are to avoid false-positive fracture diagnosis. Some QM methods that compare ratios to population reference data may classify short vertebral height correctly as non-fracture. In addition, the anterior vertebral wedge angle has been shown to increase concurrently with age-related degenerative change. Degenerative signs include degenerative disc disease, osteophytes, and endplate irregularities (Fig 5a). The ABQ method incorporates additional differential diagnoses including previous (e.g. during childhood) fractures, metabolic diseases (e.g. osteomalacia), and developmental anomalies, including anterior step deformity (depressions in the anterior portion of the vertebral endplate) in thoracic vertebrae, balloon disc, or cupid’s bow with deep inferior endplates in the lumbar vertebrae [35].

A frequent condition that resembles vertebral fractures is Scheuermann’s disease. Radiographic criteria of Scheuermann’s disease are a thoracic kyphosis greater than 45° and at least three adjacent wedge-shaped vertebral bodies of 5° or more [60, 61]. Vertebral wedging is frequently associated with endplate irregularity and Schmorl’s nodes. Elongated vertebrae and disc space narrowing can also be found in Scheuermann’s disease (Fig 5b). This vertebral wedging may be mistaken for mild vertebral fractures by QM or SQ, and Schmorl’s nodes may mimic

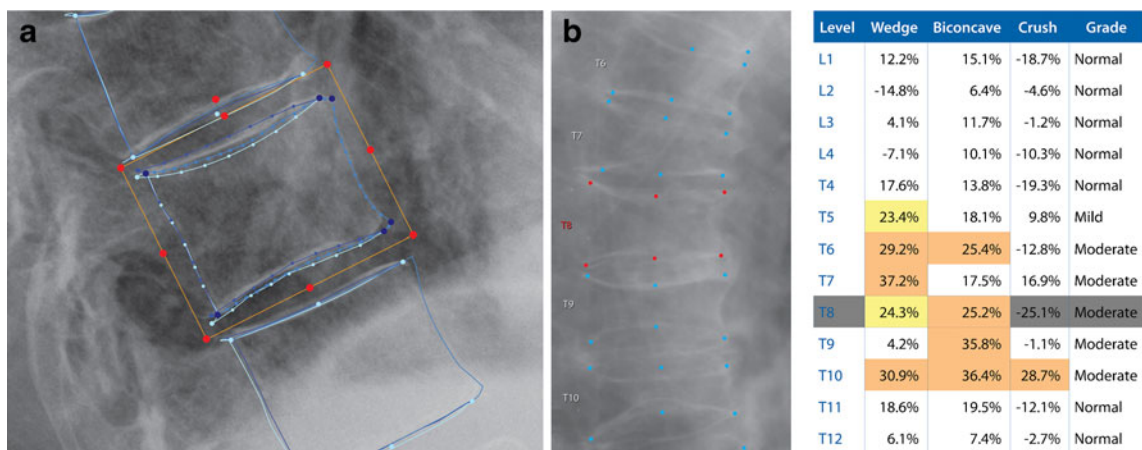


Fig. 4 Image analysis software (examinations with SpineAnalyzer®). **a** Automated contour detection. **b** Automated analysis of shape abnormality

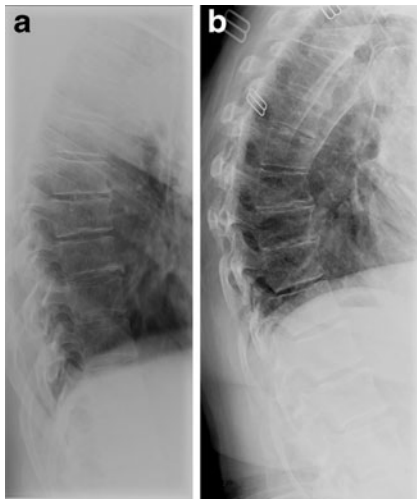


Fig. 5 Conditions that mimic vertebral fractures. **a** Degenerative changes. Very mild anterior vertebral wedging of two mid-thoracic vertebral bodies is seen along with mild spondylotic changes at the anterior vertebral margins. Note that the endplates are intact and only show mild degenerative irregularities. **b** Scheuermann's disease. In addition to marked endplate irregularity, mild anterior wedging of multiple midthoracic vertebrae is seen, resulting in increased thoracic kyphosis

endplate depression. Occasionally, but most importantly, osteoporotic vertebral fractures need to be distinguished from those resulting from malignant aetiologies, such as metastases (most commonly of primary breast, kidney, prostate, or lung neoplasms), multiple myeloma, or primary bone tumours [53]. In the majority of osteoporotic vertebral fracture cases, posterior margins of vertebrae maintain a straight or concave shape, whereas in malignant aetiologies the posterior margin is often convex.

Traumatic fractures should also be distinguished from the typically low-grade trauma osteoporotic fractures. Posterior height loss was regarded as posterior wedge in the McCloskey-Kanis method [2] and as crush deformity in SQ [4]. However, fractures involving the posterior vertebral part are typically attributable to malignancy or high-energy trauma [62] rather than to low trauma, which is most common in osteoporosis.

Inter- and intraobserver agreement of scoring methods

Inter- and intraobserver agreement seems to vary considerably within and between scoring methods. However, agreement is about precision of a study and may not necessarily relate to its validity. Nevertheless, there are several aspects that need to be considered when comparing methods. Point-placement in SQ and QM is said to be somewhat subjective, and hence inclined to influence fracture discrimination. This is particularly pertinent in the presence of borderline deformities. Also, reproducibility of SQ and ABQ scoring may to

some extent depend on the reader's training and experience [34, 44, 63].

Kim et al. have evaluated intra- and inter-reader agreement of a semi-automated quantitative morphometry software algorithm on lateral CT scout views [58]. They found intraclass correlation coefficients of 0.96 to 0.98 for vertebral heights, while kappa statistics were 0.59 to 0.69 for intra-reader and 0.67 for inter-reader agreement. Agreement for vertebral fracture classification was worse than agreement for height measurements. This was explained by the small variation of height measurement around fracture classification thresholds. Such clinically insignificant variation in height measurement can actually lead to two different fracture classifications in a considerable number of cases. Furthermore, kappa scores did not improve much even when the fracture definition was changed to include only moderate and severe fractures (i.e. deformity $\geq 25\%$).

Several publications have evaluated the inter- and intraobserver agreement of SQ alone and compared with QM methods [34, 36, 63–65]. Kappa statistics reported for SQ interobserver agreement ranged from 0.51 to 0.80 and from 0.76 to 0.93 for intraobserver agreement, respectively. The kappas for agreement between SQ and several QM methods have been reported to be lower, ranging from 0.23 to 0.59, with some improvement when fracture definition included only moderate and severe fractures (i.e. deformity $\geq 25\%$). Obviously, the agreement between different QM methods will depend on the fracture threshold chosen. Recently, semi-automated QM reading using Genant's criteria by a non-radiologist was compared with conventional SQ grading performed by experienced radiologists, finding a kappa for agreement of 0.78 [66].

Ferrar et al. have examined interobserver agreement for ABQ diagnosis of prevalent vertebral fracture in approximately 200 elderly women, finding kappa statistics of 0.74 for inter-reader agreement [63]. In general, the ABQ method has displayed low to moderate concordance with other methods. Jiang et al. found kappa statistics between 0.39 and 0.64 comparing ABQ with the QM methods developed by Eastell-Melton and McCloskey [36]. Also, ABQ has been compared with SQ observing kappa statistics of 0.30 to 0.58 [67].

Influence of scoring methods on vertebral fracture prevalence and incidence

All methods assess osteoporotic vertebral fractures with different criteria, which results in different estimates of the prevalence of the disease [36, 68]. For example, QM and SQ would not diagnose vertebral fractures in the case of endplate depression without reduced vertebral height (Fig 6a). Conversely, ABQ would not diagnose a QM-based vertebral fracture with reduced height but intact endplates (Fig 6b). In general, SQ would yield a higher number of fractures than

when applying QM, asserting that SQ would be more sensitive particularly for the detection of mild deformities [68]. However, Melton et al. have demonstrated that depending on the morphometric definition used, the prevalence of vertebral fractures ranged from 3 to 90% in their study [69]. Of all the methods, the ABQ reading results in the lowest estimations of vertebral fracture prevalence. The question remains if the higher estimates from other methods are actually due to false-positive classification of non-fracture deformities [36].

Research implications

Misclassification of vertebral fractures may result in non-differential information bias, leading to dilution of observed effects. As a consequence, true associations in studies may go undetected. This ascertainment bias can occur both in case control studies (including clinical trials investigating drug effects) and in observational studies. Therefore, scoring methods should procure the optimal classification of true vertebral fractures.

Large-scale application of standardised scoring can be difficult, with purely morphometric approaches being laborious, while the other methods will require thorough training of observers. In very large studies, especially population-based studies with an expected low prevalence of vertebral fracture cases, a technician triage system may reduce the work burden of scoring thousands of radiographs [65, 68, 70]. First, trained research technicians can triage radiographs as definite vertebral fracture, uncertain fracture, or definite normal. Finally, an expert reader may review the difficult cases and confirm vertebral fractures. Also, a step-wise evaluation process combining morphometry and

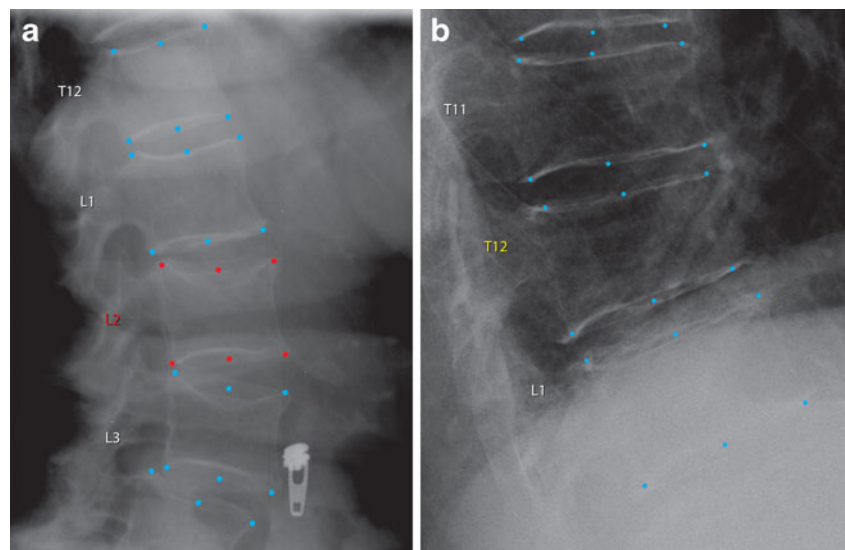
qualitative assessment represents a possible procedure to achieve a final diagnosis of vertebral osteoporosis [71].

To date, there have been few large-scale comparisons of vertebral fracture assessment methods. We are currently applying both ABQ and software-assisted QM methods to radiographs from the Rotterdam Study (all image examples included in the present article originate from this study). This study is a prospective population-based cohort that has been studying disease and disability in more than 15,000 individuals aged 45 and over since 1990 [72]. Within the ongoing research program, radiographs of approximately 11,000 participants are available, with a follow-up duration of maximally 15 years. An aim of the study is to compare the methods applied for identifying vertebral fractures. In addition, data on numerous outcomes and risk factors are available, including a comprehensive assessment of clinical fractures, BMD, and genetic determinations.

Clinical implications

It is estimated that only about one third of all vertebral fractures come to clinical attention [23]. However, assessment of vertebral fracture status, in addition to BMD, provides practical and relevant clinical information to aid the prediction of subsequent fracture risk [73]. Symptomatic and non-symptomatic vertebral fractures are both associated with decreased quality of life [2–7] and increased mortality risk [8, 9]. In the case of vertebral fracture, pharmacologic therapy is considered necessary to prevent the occurrence of future osteoporotic fractures [74]. However, as all interventions have costs and potential side effects, correct assessment of vertebral fractures is of utmost importance. Over- and underdiagnosis can have major consequences, particularly at the population level. Misdiagnosis of osteoporotic

Fig. 6 Discrepancy between scoring methods. **a** Vertebral fractures diagnosed with ABQ L2 and L3 based on endplate depression, classified as normal with QM because of normal vertebral height. **b** Vertebral fracture of T12 diagnosed with QM, based on height reduction, classified as normal with ABQ because of intact endplates



vertebral fractures will result in under- or overtreatment of patients and subsequently unnecessary costs, increased morbidity, and higher mortality.

Current definitions used by vertebral fracture scoring methods seem to be based on arbitrary cutoffs. At most, some QM methods have been established by deriving standard deviations from measurements in a sample of healthy individuals, but variation from the mean is not necessarily abnormal. The classifications show association with osteoporosis-related outcomes such as BMD and the risk of future non-vertebral and new vertebral fractures [17, 36, 65, 70, 75]. Yet, from a more clinically oriented perspective the definition of vertebral fractures should be based on cutoffs that were defined based on their ability to predict relevant outcomes, such as future osteoporotic fractures. This will require the optimal combination of true- and false-positive ratios that yield the greatest expected utility for the patient at acceptable costs to society. For optimal appraisal of future osteoporotic fracture risk it might prove necessary to refine currently available vertebral fracture scoring after comprehensive comparative studies and integrate more quantitative information that can be derived from imaging, for example three-dimensional reconstruction of vertebral shape, BMD, and measurements able to appreciate the integrity of the endplates and microarchitecture.

In conclusion, standardised and accurate scoring methods for osteoporotic vertebral fractures are desirable. There are several radiological scoring methods for osteoporotic vertebral fractures, which can be characterised as quantitative, qualitative, or semi-quantitative. Also, these standardised scoring methods can be implemented for different imaging modalities. The scoring methods each use different definitions for the diagnosis of vertebral fracture and the classification of severity. Such differences have implications for patient care and scientific research. Accurate diagnosis of vertebral fractures and differentiation from non-fracture deformities is an important aspect that depends on the expertise of the reader. Future evaluation of the concordance between methods will allow establishing their benefits and limitations, and most importantly, optimise their effectiveness for application in clinical and research scenarios.

Acknowledgements Part of the content of this article was presented as a scientific poster presentation at the American Society for Bone and Mineral Research Annual Meeting in San Diego, CA, USA, from 16–20 September 2011 and as an educational exhibit at the RSNA Annual Meeting in Chicago, IL, USA, from 27 November–2 December 2011.

References

1. Szulc P, Bouxsein ML (2011) Overview of osteoporosis: Epidemiology and clinical management. Vertebral Fracture Initiative Resource Document PART I
2. Burger H, Van Daele PL, Grashuis K et al (1997) Vertebral deformities and functional impairment in men and women. *J Bone Miner Res* 12:152–157
3. Gold DT (1996) The clinical impact of vertebral fractures: quality of life in women with osteoporosis. *Bone* 18:185S–189S
4. Nevitt MC, Ettinger B, Black DM et al (1998) The association of radiographically detected vertebral fractures with back pain and function: a prospective study. *Ann Intern Med* 128:793–800
5. Oleksik A, Lips P, Dawson A et al (2000) Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *J Bone Miner Res* 15:1384–1392
6. Oleksik AM, Ewing S, Shen W, van Schoor NM, Lips P (2005) Impact of incident vertebral fractures on health related quality of life (HRQOL) in postmenopausal women with prevalent vertebral fractures. *Osteoporos Int* 16:861–870
7. Ross PD (1997) Clinical consequences of vertebral fractures. *Am J Med* 103:30S–42S, discussion 42S–43S
8. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR (2009) Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. *JAMA* 301:513–521
9. Ensrud KE, Thompson DE, Cauley JA et al (2000) Prevalent vertebral deformities predict mortality and hospitalization in older women with low bone mass. *J Am Geriatr Soc* 48:241–249
10. Ström O, Borgström F, Kanis JA et al (2011) Osteoporosis: burden, health care provision and opportunities in the EU. *Arch Osteoporos*. doi:10.1007/s11657-011-0060-1
11. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A (2007) Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *J Bone Miner Res* 22:465–475
12. Ralston SH, Uitterlinden AG (2010) Genetics of osteoporosis. *Endocr Rev* 31:629–662
13. van der Klift M, de Laet CE, McCloskey EV et al (2004) Risk factors for incident vertebral fractures in men and women: the Rotterdam Study. *J Bone Miner Res* 19:1172–1180
14. WHO (1994) Assessment of fracture risk and its implication to screening for postmenopausal osteoporosis: Technical report series 843. World Health Organization, Geneva
15. Kelsey JL, Samelson EJ (2009) Variation in risk factors for fractures at different sites. *Curr Osteoporos Rep* 7:127–133
16. Cooper C, Atkinson EJ, O’Fallon WM, Melton LJ 3rd (1992) Incidence of clinically diagnosed vertebral fractures: a population-based study in Rochester, Minnesota, 1985–1989. *J Bone Miner Res* 7:221–227
17. Black DM, Arden NK, Palermo L, Pearson J, Cummings SR (1999) Prevalent vertebral deformities predict hip fractures and new vertebral deformities but not wrist fractures. Study of osteoporotic fractures research group. *J Bone Miner Res* 14:821–828
18. McCloskey EV, Vasireddy S, Threlkeld J et al (2008) Vertebral fracture assessment (VFA) with a densitometer predicts future fractures in elderly women unselected for osteoporosis. *J Bone Miner Res* 23:1561–1568
19. Melton LJ 3rd, Atkinson EJ, Cooper C, O’Fallon WM, Riggs BL (1999) Vertebral fractures predict subsequent fractures. *Osteoporos Int* 10:214–221
20. Lindsay R, Silverman SL, Cooper C et al (2001) Risk of new vertebral fracture in the year following a fracture. *JAMA* 285:320–323
21. Williams AL, Al-Busaidi A, Sparrow PJ, Adams JE, Whitehouse RW (2009) Under-reporting of osteoporotic vertebral fractures on computed tomography. *Eur J Radiol* 69:179–183
22. Lewiecki EM, Laster AJ (2006) Clinical review: clinical applications of vertebral fracture assessment by dual-energy x-ray absorptiometry. *J Clin Endocrinol Metab* 91:4215–4222
23. Cooper C, Shah S, Hand DJ et al (1991) Screening for vertebral osteoporosis using individual risk factors. The multicentre vertebral fracture study group. *Osteoporos Int* 2:48–53

24. Delmas PD, van de Langerijt L, Watts NB et al (2005) Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. *J Bone Miner Res* 20:557–563
25. Gehlbach SH, Bigelow C, Heimisdottir M, May S, Walker M, Kirkwood JR (2000) Recognition of vertebral fracture in a clinical setting. *Osteoporos Int* 11:577–582
26. Bartalena T, Giannelli G, Rinaldi MF et al (2009) Prevalence of thoracolumbar vertebral fractures on multidetector CT: underreporting by radiologists. *Eur J Radiol* 69:555–559
27. Kim N, Rowe BH, Raymond G et al (2004) Underreporting of vertebral fractures on routine chest radiography. *AJR Am J Roentgenol* 182:297–300
28. Guglielmi G, Diacinti D, van Kuijk C et al (2008) Vertebral morphometry: current methods and recent advances. *Eur Radiol* 18:1484–1496
29. Almen A, Tingberg A, Besjakov J, Mattsson S (2004) The use of reference image criteria in X-ray diagnostics: an application for the optimisation of lumbar spine radiographs. *Eur Radiol* 14:1561–1567
30. Hurxthal LM (1968) Measurement of anterior vertebral compressions and biconcave vertebrae. *Am J Roentgenol Radium Ther Nucl Med* 103:635–644
31. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ 3rd (1991) Classification of vertebral fractures. *J Bone Miner Res* 6:207–215
32. McCloskey EV, Spector TD, Eyres KS et al (1993) The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* 3:138–147
33. Davies KM, Recker RR, Heaney RP (1989) Normal vertebral dimensions and normal variation in serial measurements of vertebrae. *J Bone Miner Res* 4:341–349
34. Genant HK, Wu CY, van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res* 8:1137–1148
35. Ferrar L, Jiang G, Adams J, Eastell R (2005) Identification of vertebral fractures: an update. *Osteoporos Int* 16:717–728
36. Jiang G, Eastell R, Barrington NA, Ferrar L (2004) Comparison of methods for the visual identification of prevalent vertebral fracture in osteoporosis. *Osteoporos Int* 15:887–896
37. Marshall D, Johnell O, Wedel H (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 312:1254–1259
38. 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
39. The European Prospective Osteoporosis (EPOS) Study Group (2002) The relationship between bone density and incident vertebral fracture in men and women. *J Bone Miner Res* 17:2214–2221
40. Lorente-Ramos R, Azpeitia-Arman J, Munoz-Hernandez A, Garcia-Gomez JM, Diez-Martinez P, Grande-Barez M (2011) Dual-energy x-ray absorptiometry in the diagnosis of osteoporosis: a practical guide. *AJR Am J Roentgenol* 196:897–904
41. Kanis JA, Oden A, Johansson H, Borgstrom F, Strom O, McCloskey E (2009) FRAX and its applications to clinical practice. *Bone* 44:734–743
42. Cauley JA, Hochberg MC, Lui LY et al (2007) Long-term risk of incident vertebral fractures. *JAMA* 298:2761–2767
43. Kolta S, Quiligotti S, Ruyssen-Witrand A et al (2008) In vivo 3D reconstruction of human vertebrae with the three-dimensional X-ray absorptiometry (3D-XA) method. *Osteoporos Int* 19:185–192
44. Buehring B, Krueger D, Checovich M et al (2010) Vertebral fracture assessment: impact of instrument and reader. *Osteoporos Int* 21:487–494
45. Schousboe JT, Debold CR (2006) Reliability and accuracy of vertebral fracture assessment with densitometry compared to radiography in clinical practice. *Osteoporos Int* 17:281–289
46. Damilakis J, Adams JE, Guglielmi G, Link TM (2010) Radiation exposure in X-ray-based imaging techniques used in osteoporosis. *Eur Radiol* 20:2707–2714
47. Samelson EJ, Christiansen BA, Demissie S et al (2011) Reliability of vertebral fracture assessment using multidetector CT lateral scout views: the Framingham Osteoporosis Study. *Osteoporos Int* 22:1123–1131
48. Burghardt AJ, Link TM, Majumdar S (2011) High-resolution computed tomography for clinical imaging of bone microarchitecture. *Clin Orthop Relat Res* 469:2179–2193
49. Krug R, Burghardt AJ, Majumdar S, Link TM (2010) High-resolution imaging techniques for the assessment of osteoporosis. *Radiol Clin North Am* 48:601–621
50. Link TM (2012) Osteoporosis imaging: state of the art and advanced imaging. *Radiology* 263:3–17
51. Van Goethem JW, Maes M, Ozsarlak O, van den Hauwe L, Parizel PM (2005) Imaging in spinal trauma. *Eur Radiol* 15:582–590
52. Wilmink JT (1999) MR imaging of the spine: trauma and degenerative disease. *Eur Radiol* 9:1259–1266
53. Drevelgas A, Choumouzi D, Boulogianni G, Sofroniadis I (2003) Imaging of primary bone tumors of the spine. *Eur Radiol* 13:1859–1871
54. Erlemann R (2006) Imaging and differential diagnosis of primary bone tumors and tumor-like lesions of the spine. *Eur J Radiol* 58:48–67
55. Brett A, Miller CG, Hayes CW et al (2009) Development of a clinical workflow tool to enhance the detection of vertebral fractures: accuracy and precision evaluation. *Spine (Phila Pa 1976)* 34:2437–2443
56. Guglielmi G, Palmieri F, Placentino MG, D'Errico F, Stoppino LP (2009) Assessment of osteoporotic vertebral fractures using specialized workflow software for 6-point morphometry. *Eur J Radiol* 70:142–148
57. Roberts MG, Oh T, Pacheco EM, Mohankumar R, Cootes TF, Adams JE (2012) Semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. *Osteoporos Int* 23:655–664
58. Kim YM, Demissie S, Eisenberg R, Samelson EJ, Kiel DP, Bouxsein ML (2011) Intra- and inter-reader reliability of semi-automated quantitative morphometry measurements and vertebral fracture assessment using lateral scout views from computed tomography. *Osteoporos Int* 22:2677–2688
59. Link TM, Guglielmi G, van Kuijk C, Adams JE (2005) Radiologic assessment of osteoporotic vertebral fractures: diagnostic and prognostic implications. *Eur Radiol* 15:1521–1532
60. Sorensen (1964) Scheuermann's Juvenile Kyphosis: Clinical Appearances, Radiography, Aetiology, and Prognosis. Munksgaard
61. Sachs B, Bradford D, Winter R, Lonstein J, Moe J, Willson S (1987) Scheuermann kyphosis. Follow-up of Milwaukee-brace treatment. *J Bone Joint Surg Am* 69:50–57
62. Bensch FV, Kiuru MJ, Koivikko MP, Koskinen SK (2004) Spine fractures in falling accidents: analysis of multidetector CT findings. *Eur Radiol* 14:618–624
63. Ferrar L, Jiang G, Schousboe JT, DeBold CR, Eastell R (2008) Algorithm-based qualitative and semiquantitative identification of prevalent vertebral fracture: agreement between different readers, imaging modalities, and diagnostic approaches. *J Bone Miner Res* 23:417–424
64. Wu CY, Li J, Jergas M, Genant HK (1995) Comparison of semi-quantitative and quantitative techniques for the assessment of prevalent and incident vertebral fractures. *Osteoporos Int* 5:354–370
65. Black DM, Palermo L, Nevitt MC et al (1995) Comparison of methods for defining prevalent vertebral deformities: the Study of Osteoporotic Fractures. *J Bone Miner Res* 10:890–902
66. Kim YM, Demissie S, Genant HK et al (2012) Identification of prevalent vertebral fractures using CT lateral scout views: a

- comparison of semi-automated quantitative vertebral morphometry and radiologist semi-quantitative grading. *Osteoporos Int* 23:1007–1016
67. Ferrar L, Jiang G, Cawthon PM et al (2007) Identification of vertebral fracture and non-osteoporotic short vertebral height in men: the MrOS study. *J Bone Miner Res* 22:1434–1441
68. Genant HK, Jergas M, Palermo L et al (1996) Comparison of semiquantitative visual and quantitative morphometric assessment of prevalent and incident vertebral fractures in osteoporosis The Study of Osteoporotic Fractures Research Group. *J Bone Miner Res* 11:984–996
69. Melton LJ 3rd, Wenger DE, Atkinson EJ et al (2006) Influence of baseline deformity definition on subsequent vertebral fracture risk in postmenopausal women. *Osteoporos Int* 17:978–985
70. Black DM, Palermo L, Nevitt MC, Genant HK, Christensen L, Cummings SR (1999) Defining incident vertebral deformity: a prospective comparison of several approaches. The Study of Osteoporotic Fractures Research Group. *J Bone Miner Res* 14:90–101
71. Ziegler R, Scheidt-Nave C, Leidig-Bruckner G (1996) What is a vertebral fracture? *Bone* 18:169S–177S
72. Hofman A, van Duijn CM, Franco OH et al (2011) The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 26:657–686
73. Siris ES, Genant HK, Laster AJ, Chen P, Misurski DA, Krege JH (2007) Enhanced prediction of fracture risk combining vertebral fracture status and BMD. *Osteoporos Int* 18:761–770
74. Chen JS, Sambrook PN (2011) Antiresorptive therapies for osteoporosis: a clinical overview. *Nat Rev Endocrinol* 8:81–91
75. Finigan J, Greenfield DM, Blumsohn A et al (2008) Risk factors for vertebral and nonvertebral fracture over 10 years: a population-based study in women. *J Bone Miner Res* 23:75–85