Abstract  Vertebral fractures are the hallmark of osteoporosis, and occur with a higher incidence earlier in life than any other type of osteoporotic fractures. It has been shown that both symptomatic and asymptomatic vertebral fractures are associated with increased morbidity and mortality. Morbidity associated with these fractures includes decreased physical function and social isolation, which have a significant impact on the patient’s overall quality of life. Since the majority of vertebral fractures do not come to clinical attention, radiographic diagnosis is considered to be the best way to identify and confirm the presence of osteoporotic vertebral fractures in clinical practice. Traditionally, conventional lateral radiographs of the thoracolumbar spine have been visually evaluated by radiologists or clinicians to identify vertebral fractures. The two most widely used methods to determine the severity of such fractures in clinical research are the semiquantitative assessment of vertebral deformities, which is based on visual evaluation, and the quantitative approach, which is based on different morphometric criteria. In our practice for osteoporosis evaluation we use the Genant semiquantitative approach: an accurate and reproducible method tested and applied in many clinical studies. The newest generation of fan-beam DXA systems delivering “high-resolution” lateral spine images offers a potential practical alternative to radiographs for clinical vertebral fracture analysis. The advantages of using DXA over conventional radiographic devices are its minimal radiation exposure and high-speed image acquisition. It also allows combined evaluation of vertebral fracture status and bone mass density, which could become a standard for patient evaluation in osteoporosis. The disadvantage of DXA use is that upper thoracic vertebrae cannot be evaluated in a substantial number of patients due to poor imaging quality. We truly believe that the that there is a major role for radiologists and clinicians alike to carefully assess and diagnose vertebral fractures using standardized grading schemes such as the one outlined in this review. Quantitative morphometry is useful in the context of epidemiological studies and clinical drug trials; however, the studies would be flawed if quantitative morphometry were to be performed in isolation without additional adjudication by a trained and highly experienced radiologist or clinician.

Keywords  Osteoporosis · Vertebral fractures · Semiquantitative assessment · Bone mineral density · Quantitative morphometry
Introduction

Osteoporosis is a serious public health problem. The incidence of osteoporotic fractures increases with age. As life expectancy increases for a greater proportion of the world’s population, the financial and human costs associated with osteoporotic fractures will multiply exponentially. According to the International Osteoporosis Foundation, more than 40% of middle-aged women in Europe will suffer one or more osteoporotic fracture during their remaining lifetime [23].

Vertebral fractures are the hallmark of osteoporosis and occur with a higher incidence earlier in life than any other type of osteoporotic fractures, including hip fractures [34]. The importance of fragility fractures, of which vertebral fractures are the most common, was acknowledged by the World Health Organization classification criteria for osteoporosis evaluation [51]. The criterion of the World Health Organization defines “severe osteoporosis” as “low bone mass (T score below –2.5) in the presence of one or more fragility fractures.”

The definition of osteoporosis is centered on the level of bone mass, which is measured as bone mineral density (BMD). BMD measurements are widely used to estimate the risk of osteoporotic fractures and individuals who are at risk for osteoporotic fractures are usually referred for BMD measurements under the current standard of care. In addition, many other risk factors have been identified, some of which are known to add to the risk independently of BMD measurements. The combination of BMD with such risk factors increases the gradient of risk/standard deviation than that achieved by BMD alone. Several clinical trials have demonstrated that a substantial improvement in the assessment of the risk for future fractures can be accomplished by the assessment of prevalent vertebral fractures in combination with BMD measurements [2, 5, 15, 27, 31, 36, 39, 41]. Nonetheless, it remains a common clinical practice to consider “low” BMD to be a risk factor irrespective of the presence of vertebral fractures.

Clinical identification of vertebral fractures

It has been shown that both symptomatic and asymptomatic vertebral fractures are associated with increased morbidity [9] and mortality [8, 22, 35]. Morbidity associated with these fractures includes decreased physical function and social isolation, which have a significant impact on the patient’s overall quality of life [16]. Still, it remains difficult to determine the exact incidence of osteoporotic vertebral fractures that occur annually, as a substantial proportion remains clinically undetected. Large-scale prospective studies demonstrate that only about one of four vertebral fractures becomes clinically recognized [7]. This is due to both the absence of specific symptoms in some and the difficulty in determining the cause of possible physical symptoms such as pain or height loss. Therefore the evaluation of spinal radiographs for prevalent and incident vertebral fractures is important in both clinical and epidemiological evaluation of patients with established osteoporosis and populations at risk for developing it. Fewer than 1% of back pain episodes are related to vertebral fractures [10]. Therefore vertebral fractures are often not suspected in patients reporting back pain, unless associated with trauma. Trauma-related fractures, however, are not considered as classical osteoporotic fractures. Historical height loss is also difficult to assess clinically. Some height loss is expected with aging due to compression of the intervertebral discs and postural changes. However, height loss could also be due to multiple fractures, which represent significant and irreparable damage. Therefore it has been concluded that height loss is an unreliable indicator of fracture status until it exceeds 4 cm [9]. As a result vertebral fractures are often not being considered in clinical patient evaluation, and it is relatively uncommon for patients to be referred for radiographs in the course of osteoporosis testing. Improvements in detecting and reporting vertebral fractures in patients with osteoporosis would increase the potential of therapeutic intervention to prevent subsequent fractures.

Radiographic assessment of vertebral fractures

Radiographic diagnosis is considered to be the best way to identify and confirm the presence of osteoporotic vertebral fractures in clinical practice. Traditionally, conventional lateral radiographs of the thoracolumbar spine have been visually evaluated by radiologists or clinicians to identify vertebral fractures. However, there is still no internationally agreed definition for vertebral fracture. One global prospective study (the IMPACT study [6]) compared the results of local radiographic reports from five continents with that of subsequent central readings in more than 2,000 postmenopausal women with osteoporosis. This study demonstrated that vertebral fractures were frequently underdiagnosed radiologically worldwide, with false-negative rates as high as 30% despite a strict radiographic protocol that provided an unambiguous vertebral fracture definition and minimized the influence of inadequate film quality. It was concluded that the failure was a global problem attributable to either lack of radiographic detection or use of ambiguous terminology in reports. Therefore it is very important to use standardized methods for the visual assessment of vertebral fractures.

Several standardized approaches to describe vertebral fractures have been proposed. They may serve to facilitate the diagnosis of osteoporosis and to assess the severity or progression of the disease as well as to rule out nonfracture deformities or normal variants.
Visual semiquantitative methods of vertebral fracture assessment

The first standardized approach was introduced by Smith et al. [44] in 1960. They introduced a classification of vertebral deformities as diagnosed from lateral thoracolumbar radiographs for the purpose of diagnosing the severity of osteoporosis. This method grades only the most severely deformed vertebra on the radiograph. In 1968 Meunier [33] proposed an approach in which each vertebra is graded according to its shape or deformity. Grade 1 is assigned to a normal vertebra that has no deformity, grade 2 to a biconcave vertebra, and grade 4 to an endplate fracture or a wedged or crushed vertebra. Using this approach vertebral bodies T3 (or T7) to L4 are evaluated. A “radiological vertebral index” can be calculated as the sum of the grades of all vertebrae, or as the quotient of this sum and the number of the vertebrae.

Kleerkoper et al. [26] modified Meunier’s radiological vertebral index and introduced the so-called “vertebral deformity score.” In the vertebral deformity score each vertebra from T4 to L5 is assigned an individual score from 0 to 3 depending on the type of deformity. This grading scheme is based on the reduction in the anterior, middle, and posterior vertebral heights ($h_a$, $h_m$, and $h_p$, respectively).

A vertebral deformity (graded 1–3) is present when $h_a$, $h_m$, or $h_p$ is reduced by at least 4 mm or 15%. This score, as with Meunier’s radiological vertebral index, still relies very much on the type of deformity, i.e., the vertebral shape, and there would have to be changes in vertebral shape in order to account for incident vertebral fractures on follow-up radiographs. Furthermore, the majority of vertebral fractures consist of a combination of wedge and endplate deformities, and less frequently posterior deformities. Therefore an examiner’s distinction among these deformities is often quite subjective.

A vertebral deformity does not always represent a vertebral fracture, but a vertebral fracture is always a vertebral deformity. From a radiological prospective, there are many potential differential diagnoses for vertebral deformities – osteoporotic fracture, posttraumatic deformity, degenerative remodeling, Scheuermann’s disease (juvenile kyphosis), congenital anomaly, neoplastic deformity, and Paget’s disease – and the correct qualitative classification of vertebral deformities can be accomplished only by visual inspection and expert interpretation of the radiograph. This perspective on vertebral fracture diagnosis is perhaps reflected at its best in the semiquantitative fracture assessment method proposed by Genant et al. [12, 13, 14, 15]. This method provides an insight into the severity of a

Fig. 1 Schematic diagram of semiquantitative grading scale for vertebral fractures. (From Genant et al. [13])

Normal (Grade 0)

Wedge deformity

Biconcave deformity

Mild fracture (Grade 1)

Moderate fracture (Grade 2)

Severe fracture (Grade 3)
fracture which is assessed solely by visual estimation of the extent of a vertebral height reduction and morphological change, and vertebral fractures are differentiated from other, nonfracture deformities. In Genant’s visual semiquantitative assessment (Fig. 1) each vertebra receives a severity grade based upon the visually apparent degree of vertebral height loss. Unlike the other approaches the type of the deformity (wedge, biconcavity, or compression) is no longer linked to the grading of a fracture in this approach.

Thoracic and lumbar vertebrae from T4 to L4 are graded on visual inspection and without direct vertebral measurement as normal (grade 0), mildly deformed (grade 1: reduction of 20–25% of height and 10–20% of projected vertebral area), moderately deformed (grade 2: reduction of 26–40% of height and 21–40% of projected vertebral area), and severely deformed (grade 3: reduction of >40% of height and projected vertebral area; Fig. 2). A grade 0.5 designates “borderline” vertebrae that show some deformation but cannot be clearly assigned to grade 1 fractures is sometimes also utilized. In addition to height reductions, careful attention is given to alterations in the shape and configuration of the vertebrae relative to adjacent vertebrae and expected normal appearances. These features add a strong qualitative aspect to the interpretation. For example, vertebral deformities due to degenerative changes should be ruled out, whereas an endplate vertebral fracture can be identified without a 20% reduction in the vertebral height. Nevertheless, in experienced, highly trained hands, it makes the approach both sensitive and specific. A “spinal fracture index”) can be calculated from this semiquantitative assessment as the sum of all grades assigned to the vertebrae divided by the number of the evaluated vertebrae.

An advantage of this semiquantitative approach over other standardized visual approaches is that the severity of the deformation as the reduction in vertebral height means can be assessed from serial films and is especially useful for the interpretation of incident fractures. It considers the continuous character of vertebral fractures and makes a meaningful interpretation of follow-up radiographs possible. Furthermore, inevitably arbitrary decisions regarding wedge, endplate, or crush deformities, as assessed in some grading schemes, are not necessary since most fractures

Fig. 2 Lateral thoracic radiograph shows a grade 3 fracture T8 and grade 2 fractures of T9 and T11

Fig. 3 Degenerative remodeling in middle-thoracic region simulating wedge deformities
contain a combination of these features, influenced by the local biomechanics of the spinal level.

The Genant’s semiquantitative method has been tested and applied in a number of clinical drug trials and epidemiological studies [15, 20, 47, 50, 52]. The reproducibility of the method for the diagnosis of prevalent and incident vertebral fractures was found to be high, with intraobserver agreement of 93–99% and interobserver agreement of 90–99%. This indicates that close agreement among readers can be reached using this standardized visual semiquantitative grading method, and that subjectivity in the readings can be reduced. This accounts for experienced and relatively inexperienced readers with reasonable results.

There are limitations of this semiquantitative grading scheme that may also apply to other standardized approaches. For example, from the morphometric data on normal subjects we know that vertebrae in the middle thoracic spine (especially in women) and thoracolumbar junction (especially in men) are slightly more wedged than in other regions (Fig. 3) [3, 30, 32, 40]. As a result these normal variations may be misinterpreted as mild vertebral deformities, thereby falsely increasing prevalence values for vertebral fractures. The same applies to a lesser extent to the middle to lower lumbar spine, where some degree of biconcavity is frequently observed [26, 45]. Accurate diagnosis of prevalent fractures which requires distinguishing between normal variations and the degenerative changes from true fractures still depends on the experience of the observer. It has been argued that the diagnosis of mild vertebral fractures (grade 0.5–1) in particular may be quite subjective, and that these fractures may be unrelated to osteoporosis [45]. However, mild fractures are also associated with a lower bone density and to a certain extend predict future vertebral fractures [1].

Other limitations may apply for the diagnosis of incident fractures. The reader may sometimes feel that even though a further height reduction is seen in a previous vertebral fracture, it may not be justified to assign a higher fracture grade on a serial radiograph, since some degree of settling or remodeling generally occurs. Therefore in general, serial radiographs including the baseline radiograph of a patient should be viewed together so that incident fractures can be readily identified as only those progressive changes that lead to a full increase in deformity grade or from a questionable deformity (grade 0.5) to a definite fracture.

**Quantitative morphometry and its comparison with the semiquantitative methods**

Quantitative morphometric assessment of vertebral deformity was introduced in order to obtain an objective and reproducible measurement, using rigorously defined point placement and well-defined algorithms for fracture definition [3, 42]. Typically six points are used to derive the anterior height ($h_a$), the central (middle or middle-vertebral, $h_m$) height, and the posterior height ($h_p$, Fig. 4). This exclusively quantitative approach has, however, a number of drawbacks including projectional effects that significantly influence the reliability of these measures performed in isolation.

In general, a substantial number of mild deformities detected by visual reading are missed by the quantitative technique when applying the common threshold values for reduction in vertebral heights such as 15–20% or 3 SD decrease. Furthermore, a significant number of false positives are found with quantitative techniques. The choice of point placement in the quantitative technique, but especially the choice of the threshold for defining vertebral deformity, gives results that vary in specificity and sensitivity. Most of the moderate to severe deformities are detected by both techniques. However, only expert visual evaluation can detect mild and subtle deformities, as well as appreciate anatomical, pathological and technical issues that bear on the evaluation of fracture detection.

The strength of a semiquantitative approach is that it makes use of the entire spectrum of visible features that are helpful in identifying deformities [15, 49]. The visual interpretation, when performed by the expert eye, also separates true deformities from normal or anomalous vertebrae. In addition to changes in dimension, vertebral deformities are generally detected visually by the presence of endplate deformities, the lack of parallelism of the endplates, and the general altered appearance compared with...
neighboring vertebrae. Some of these visual characteristics are not captured by the six-digitization points used in quantitative techniques; this can cause some deformities to remain undetected. For example, only an experienced observer can make the subtle distinctions between a fractured endplate and wedge shaped appearance caused by the remodeling of the vertebral bodies in degenerative disc disease (Fig. 3). This is often interpreted as a wedge fracture in quantitative studies.

In the absence of distinct characteristics of a fracture, however, a reader using a visual approach could rather arbitrarily consider a mild wedge deformity normal, anomalous, or fractured; in such a case, a well-defined quantitative criterion could be useful. Even here, however, with borderline wedge deformity, small subjective differences in joint placement could result in considerable variation in fracture/nonfracture discrimination of sequential films or even on the same film.

Most incident fractures, as with prevalent fractures, are easily identifiable visually on sequential radiographs. The unavoidable variation in position and parallax may result in differences in point placement on follow-up radiographs. This can result in the morphometric detection of an incident fracture that would be interpreted visually as simply an alteration in projection. These sources of false-positive or false-negative interpretation are especially common when parallax problems due to radiographic technique or patient positioning are encountered.

Intraobserver variability for a semi-quantitative approach depends on experience and training. The same however, is true for digitizing techniques: an experienced observer is more consistent in the placement of the points for digitization.

A number of comparative studies have evaluated the relative performance of the quantitative morphometric and the semiquantitative methods and moderate correlations were found in most of them [1, 17, 29, 52]. The concordance was high for fractures defined as moderate or severe by semiquantitative reading. There was, however, a significant discordance for fractures defined as mild in the semiquantitative reading. Additionally, the interobserver agreement was demonstrated to be better for the visual semiquantitative approach. The authors of these studies concluded that quantitative morphometry should not be performed in isolation, particularly when applying highly sensitive morphometric criteria at low threshold levels without visual assessment to confirm the detected prevalent or incident vertebral deformities as probable fractures.

**Standardization of visual approaches to vertebral fracture assessment**

In an effort to develop a standardized consensus protocol for the visual assessment of vertebral fractures, the United States National Osteoporosis Foundation’s Working Group on Vertebral Fractures suggested the following procedural requirements for a qualitative (semiquantitative) assessment of vertebral fractures in osteoporosis research [25]:

- Assessments should be performed by a radiologist or trained clinician who has specific expertise in the radiology of osteoporosis.
- Qualitative and semiquantitative assessments should be performed according to a written protocol of fracture definitions, which are sufficiently detailed that the readings can be reproduced by other experts. Reference to an atlas of standard films or illustrations may be helpful. It is recommended that a standardized protocol be developed by a consensus of expert radiologists.
- The definition of fracture should include deformities of the endplates and anterior borders of vertebral bodies, as well as generalized collapse of a vertebral body.
- Grading of the extent of each fracture should employ discrete, mutually exclusive categories. An atlas of standard films and illustrations may again help to assure consistency.

There is some subjectivity in each method, and performing the grading in discrete, exclusive categories may be problematic at times, particularly for prevalent fractures. However, for the assessment of vertebral fractures in the form of a fracture/nonfracture dichotomy, trained readers have achieved excellent results. After all, the fracture/nonfracture distinction may be the most important, and the semiquantitative standardized grading schemes may be the instruments to make this diagnosis reliable and valid.

Ensuring the reliability of the interpretation of incident vertebral fractures on serial radiographs requires close attention to the procedure. Serial radiographs of a patient should always be viewed together in chronological order to accomplish a thorough and reliable analysis of all new fractures. Because a vertebral fracture is a permanent event that is unlikely to vanish on follow-up radiograph, temporal blinding does not appear to be any use: most readers easily identify a temporal sequence of films by new deformities as well as by progressive disc degeneration and osteophyte formation, which are universal among the elderly.

**Alternatives to radiographic assessment of vertebral fractures**

Because of the difficulty in identifying vertebral fractures clinically, and the practical difficulties preventing routine radiographic assessment at the point of care, vertebral fracture status is frequently unknown at the time of patient evaluation for BMD [18]. Hence the interest in morphometric assessment from dual X-ray absorptiometry (DXA) images was a natural consequence of the need for quantitative fracture evaluation in pharmaceutical trials. The main advantage of the morphometric X-ray absorptio-
try technique is that the radiation dose to the patient is substantially reduced compared with conventional radiography. The use of “high-resolution” lateral spine images, obtained with fan-beam X-ray bone densitometry systems (Fig. 5), offers a potential practical alternative to radiographs for clinical vertebral fracture analysis. “High-resolution” fan-beam DXA systems, utilizing technology similar to that used by computed tomography (CT) systems, can image the lateral spine in as little as 10 s. In fact CT scout scans, with about the same image resolution as fan-beam DXA scans, have been used for vertebral fracture identification [24, 43, 48].

As with radiographs, however, CT images are expensive and are not available clinically without referral. Consequently CT is not generally an option unless performed in conjunction with quantitative CT for BMD assessment. In contrast, DXA images can be performed at the point of care, in conjunction with standard BMD determination, with a radiation dose as much as 100 times lower than that of conventional radiographs. The most notable strength of radiographs, of course, is image resolution, which is superior to that of DXA images.

DXA images provide several advantages. The digital nature allows for electronic data storage, digital image enhancement and processing, as with magnification and contrast adjustment, which is not possible with conventional radiographic techniques. Cone-beam distortion, inherent in the radiographic technique, is not present when using the scanning fan-beam geometry of DXA devices. Low-dose, single-energy acquisition modes are substantially faster than dual energy scan modes due to substantially lower signal to noise in the images and can be performed during suspended respiration. High-dose, dual-energy acquisitions, while slower, generally provide higher bone contrast images and sometimes reduce artifacts.

The use of fan-beam DXA images for quantitative (morphometric) assessment of spinal fractures has been reported in both research applications and pharmaceutical trials [4, 11, 19, 21, 28, 37, 38, 46]. Clinical studies demonstrated the feasibility of visual evaluation of fan-beam lateral DXA spine images compared to conventional lateral spine radiographs in postmenopausal women, with a strong overall agreement of 96.3% [37, 38]. This agreement was approximately as strong as that found among different morphometric techniques [15, 21]. The images permitted visual assessment of about 90% of all vertebrae. The main shortcoming of the MXA scans in comparison with conventional radiographs is the inferior image quality that limits the evaluation of vertebrae in the upper thoracic spine. This is less of a concern if MXA is used as a screening tool for conventional radiography and this approach may help reduce the radiation dose in the diagnosis and monitoring of osteoporosis.

Conclusion

Vertebral fractures are the most common type of osteoporotic fracture, occurring in a substantial portion of the elderly population. Most new vertebral fractures, even painful ones, remain unrecognized by patients and their physicians. It is established that the presence of a vertebral fracture is a strong risk factor for subsequent osteoporotic fractures, and that those with low bone density and vertebral fractures are at highest risk. Large-scale clinical trials have demonstrated that osteoporosis therapies can reverse bone loss and reduce fracture rates, and that these benefits are most pronounced in patients with low BMD and vertebral fractures. Clinical guidelines promulgated by the National Osteoporosis Foundation, International Osteoporosis Foundation, and others recognize the importance of vertebral fractures, along with BMD, as the key risk factors for use in patient evaluation. However, while BMD is widely used in patient evaluation, radiological assessment of vertebral fractures is commonly not performed, or if performed, is inadequately standardized and interpreted. By understanding the clinical principles of osteoporosis diagnosis and management provided in this document and by adopting the radiological guidelines for assessing vertebral fractures provided herein, clinicians worldwide can contribute substantially to reducing the consequences of this important disease.
References

33. Meunier P (1968) La dynamique du remaniement osseux humain, etudee par lecture quantitative de la biopsie osseuse. Lyon