ORIGINAL ARTICLE

Diagnosis of osteoporosis: visual assessment on conventional versus digital radiographs

Sabine Wagner · Axel Stäbler · Harald Sittek
Harald Bonel · Gerlinde Laeverenz
Maximilian F. Reiser · Andrea Baur-Melnyk

Abstract In many radiological departments conventional radiography has been replaced by digital radiography. Therefore, the purpose of this study was to analyze the visual detection of osteopenia/osteoporosis with both digital and conventional radiographs. In 286 patients we retrospectively evaluated radiographs of the lumbar spine in two planes. One hundred twenty-eight patients had conventional and 158 patients had digital radiographs. Patients with pre-existing vertebral fractures were excluded. Four experienced musculoskeletal radiologists blinded to the values of DXA and to the patients' ages assessed independently from each other whether the bone density of the lumbar spines was normal or decreased. The results of dual X-ray absorptiometry served as the standard of reference. The threshold value for the diagnosis of osteopenia was a T-score less than −1 SD according to the WHO classification of osteoporosis. Sensitivity/specificity was 86%/36% for conventional and 72%/47% for digital radiographs. The overall diagnostic accuracy was 68% for conventional and 64% for digital radiographs. Eighty percent of the patients with osteopenia and 96% of the patients with osteoporosis were correctly assessed as true positive on conventional radiographs and 65% (osteopenia) and 82% (osteoporosis) on digital radiographs. Interobserver agreement was markedly lower for digital (35%) than for conventional radiographs (73%). However, the differences were not statistically significant. There is no major difference in diagnostic accuracy in the assessment of osteopenia/osteoporosis using digital and conventional radiographs, respectively. However, the high interobserver variance on digital radiographs indicates that visual assessment of osteoporosis/osteopenia is problematic, which may be due to image processing and postprocessing algorithms that manipulate the visual aspect of bone density.

Keywords Detection rate · Diagnostic accuracy · Digital/conventional radiography · DXA · Osteopenia · Osteoporosis

S. Wagner (✉) · H. Sittek · G. Laeverenz · M. F. Reiser
A. Baur-Melnyk
Department of Clinical Radiology, Ludwig Maximilian University,
Marchioninistrasse 15, 81377, Munich, Germany
E-mail: Wagner.Sabine@med.uni-muenchen.de
Tel.: +49-89-70953620
Fax: +49-89-70958832

A. Stäbler
Department of Radiology, Orthopedic Clinic of
Munich Harlaching, Munich, Germany

H. Bonel
Department of Radiology, Inselspital Bern,
Bern, Switzerland

Introduction

Osteoporosis is one of the most common diseases of bone, especially in elderly women. Because of a higher life expectance, the incidence of osteoporosis has even increased in the last several decades [1]. In 1993 osteoporosis was defined as a systemic bone disease that is characterized by a low bone mineral density and a loss of microarchitecture of bone structure, with increased risk of fractures [2]. As osteoporotic fractures cause significant morbidity, mortality and costs for the public health system [3, 4], the early detection of osteoporosis is highly important because effective therapeutic strategies are available to prevent vertebral fractures.

Radiographs are often the first modality that raises suspicion of osteoporosis and initiates further diagnostic work up. Various characteristic signs on radiograms are indicative of "osteoporosis". These include increased radiotranslucency of bones due to a decreased bone mass. The cancellous bone is rarefied and the cortex of the vertebrae is accentuated, which is described as the "empty box sign," and vertebrae show relatively pronounced endplates. The reduction of the cancellous bone might result in spontaneous fractures, which in the spine present as codfish or wedge-shaped vertebrae [5, 6]).

Dual X-ray absorptiometry (DXA) [7] is an established method often performed when the suspicion of osteopenia/osteoporosis has been raised on radiographs that allows for reliable diagnosis and quantification of osteoporosis [3, 8]. With DXA, the bone mineral density of the vertebral bodies and the proximal femur is measured as these have the highest risk for fracture [9, 10]. According to the WHO classification of 1994 the T-value that represents the standard deviation of the median peak bone mass compared to a young-aged patient sample is used for the diagnosis and classification of the severity of bone substance loss: standard deviations (SD) between −1 to −2.5 define "osteopenia," whereas SDs < −2.5 define "osteoporosis".

In many radiological departments digital radiography has replaced conventional radiography over the last years. Advantages of digital radiography include digital image transfer and archiving, postprocessing and dose reduction. The purpose of this study was to evaluate the visual detection rate of bone mineral loss from digital radiographs in comparison to conventional radiographs.

Patients and Methods:

A total of 500 patients examined in the years 1999 until 2003 were retrospectively evaluated. Diagnostics were performed using a conventional film/screen technique until August 2001; since then, digital radiography has been introduced. Two hundred eighty-six patients were included in this retrospective study. One hundred fifty-eight of these patients (132 female patients with a mean age of 54.5 years, 25 male patients with a mean age of 54.2 years; mean age of all patients 54 years, range 21–80 years) had digital radiographs, 128 patients (95 female patients with a mean age of 55.3 years, 33 male patients with a mean age of 48.3 years; mean age of all patients 51 years, range 18–80 years,) had conventional radiographs of the lumbar spine taken in two planes.

Patients were included in the trial when DXA measurements and high quality radiographs taken in two planes within a period of 8 weeks before or after the DXA measurement were available. The reason for the acquisition of radiographs and DXA-measurements of the lumbar spine was mainly exclusion of clinically suspected osteoporosis.

Exclusion criteria were the presence of vertebral fractures in order to avoid indirect signs of osteoporosis. Patients with osteochondrosis/sclerosis and aortic sclerosis were also excluded to prevent false positive DXA values. Moreover, patients with endocrine disorders, neoplasia, other severe systemic diseases and steroid or biphosphonate therapy were excluded since those disorders and conditions result in alterations of bone density and structure.

All digital radiographs were acquired with storage phosphor technology on an AGFA ADC Compact system (automatic exposure, anterior-posterior radiographs 75 kV, lateral radiographs 90 kV) and viewed on a Siemens Magic View workstation. Readers were allowed to use postprocessing options such as windowing and zooming. The conventional radiographs were acquired on a Philips Diagnost H (film/screen technique, automatic exposure, anterior-posterior radiographs 75 kV, lateral radiographs 90 kV) and presented to the readers on film (image format 20/30 cm).

DXA values (Hologic QDR 1000) of the lumbar spine (lumbar vertebrae 1 to 4) were used as the reference standard for the diagnosis of osteopenia/osteoporosis. According to the WHO classification of osteopenia and osteoporosis, patients were classified into three categories: patients with normal bone mineral density (T-score > −1 SD), osteopenia (T-score < −1 SD and > −2.5) and osteoporosis (T-score < −2.5 SD).

In the patient group with digital radiographs DXA showed a normal BMD in 49 patients (31%), an osteopenic BMD in 60 patients (38%) and an osteoporotic BMD in 49 patients (31%). In the patient group with conventional radiographs DXA showed a normal BMD in 46 patients (35.9%), an osteopenic BMD in 56 patients (43.8%) and an osteoporotic BMD in 26 patients (20.3%).

One hundred fifty-eight digital images and 128 conventional images were presented separately to four experienced musculoskeletal radiologists in reading sessions with a subseries of 40–50 images. Individual patient data were masked in order to avoid a reader bias. For conventional and digital images anterior-posterior and lateral views were analyzed. Each reader determined whether the patient had normal or reduced bone mineral density according to the usual criteria for osteopenia/osteoporosis (increased radiolucency of bone, rarefication of cancellous bone and pronounced vertebral endplates).

For statistical analysis sensitivity, specificity and positive and negative predictive values were determined. Mean values of all observer results were calculated, and interobserver variance was determined. An individual score was given to each patient, ranging from 0 (no correct diagnosis was given by any of the four readers) to 4 points (all four readers diagnosed the radiogram correctly). Mean values of conventional scores as well as digital scores were calculated. In order to analyze whether the differences between conventional and digital radiographs were of statistical significance, the scores were compared with the exact median test for two samples. A significance level of 0.05 was chosen.

Results

Overall diagnostic accuracy (true positive and true negative results) was 64% for digital and 68% for conventional radiographs. Sensitivity (true positives) for pathologic mineral density (T-value ≤ −1 SD) resulted in 86% for conventional versus 72% for digital

radiographs (Figs. 1 and 2). Specificity (true negatives) was 36% for conventional and 47% for digital radiographs (Figs. 1 and 3). False negative results were found in 14% on conventional and in 28% on digital radiographs (Fig. 4) and false positive answers were given in 53% on digital and 63% on conventional radiographs (Fig. 5).

In the patient group with osteopenia (DXA < -1 SD and > -2.5 SD) sensitivity (true positive findings) was 80% for conventional versus 65% for digital radiographs. Specificity (true negative findings) was 36% for conventional and 47% for digital X-rays (Fig. 6).

In the patient group with osteoporosis (DXA < -2.5 SD) sensitivities were 96% for conventional X-rays versus 82% for digital radiographs. Specificity was 36% for conventional and 47% for digital radiographs (Fig. 7). In Tables 1 and 2, the results of the readings of the conventional and digital images of all four readers and the corresponding positive and negative predictive values (ppv and npv) as well as mean values are shown.

With increasing mineral loss as assessed by DXA, an increasing number of patients was correctly diagnosed as true positive (Fig. 8). There were more true positive diagnoses for conventional radiographs when compared with digital radiographs. However, specificity was lower on conventional than on digital radiographs. False-positive evaluations occured with all T-values. Eighteen percent of digital radiographs with a T-score less than -3 SD and 6% of conventional images with a T-score below 3 SD were assessed as normal.

Interobserver agreement between all four readers, independent of whether the diagnosis was correct was 73% for conventional X-rays and 35% for digital images. Interobserver agreement in combination with a correct diagnosis was 51% for conventional and 26% for digital images.

The mean score for conventional radiographs resulted in 2.55 points, while the mean score for digital radiograms was 2.45 points. The differences between the scores for conventional and digital radiographs were not statistical significant ($P = 0.1285$, exact median test for two samples).

Discussion

The detection of osteopenia/ osteoporosis on radiograms of the spine or chest is important in order to reduce morbidity, mortality and costs due to fractures

	Conventional	Digital
Sensitivity:	86%	72%
Specificity:	36%	47%

Fig. 1 Results in patients with osteoporosis (DXA: T-score < -2.5 SD)

[11, 12]. Effective therapies, most notably bisphosphonates, are widely available. In the literature, several studies on the detection of vertebral fractures due to osteopenia/osteoporosis on X-rays have been published. Genant et al. demonstrated that the visual detection rate of vertebral fractures can be enhanced by morphometric measurements of the vertebral height [13]. In 1981 Kovarik compared various semiquantitive methods for the diagnosis of osteoporosis and correlated the results with photon absorption densitometry, which is no longer used for the quantification of bone density [14]. Only one study addressed the visual detection rate of osteopenia/osteoporosis of the spine on lateral chest radiographs. A multi-reader analysis with nine readers was performed and compared to the results of DXA measurements of the lumbar spine [15]. As of yet, studies on the visual detection rate of osteopenia/osteoporosis on digital in comparison to conventional radiographs have not been performed.

In our study on 286 patients, diagnostic accuracy (i.e., true positive and true negative findings) was 68% for conventional and 64% for digital radiographs without significant differences between digital and conventional radiography. Similarly, in the study of Jergas et al., a diagnostic accuracy between 68 and 76% was found [4]. However, it is difficult to compare these results because Jergas et al. defined osteopenia with a T-score below -2 SD in contrast to our study in which we used the revised WHO criteria of 1994 (osteopenia: T-score below -1 SD). In addition, Jergas et al. used chest X-rays that were performed with 120 kV and were therefore not appropriate for the evaluation of bony structures as spinal radiograms performed with 75–90 kV.

Specificity was low both for conventional (36%) and for digital (47%) radiographs. However, the readers achieved better results for sensitivity than for specificity, with a sensitivity of 86% for conventional versus 72% for digital radiographs. The results may have been biased by the fact that readers were aware of taking part in a study on the assessment of osteoporosis, which may have led to overdiagnosis of osteopenia and osteoporosis. With rising bone mineral density, however, false-positive evaluations decreased as expected.

When comparing conventional and digital radiographs, conventional images had higher sensitivity and a lower specificity than digital images, with more true-positive findings and a slightly better mean score in the conventional images group. We assume that a loss in mineralization content is more visible on conventional than on digital images, because conventional radiographs have a defined absolute contrast, whereas digital images are individually post-processed in order to optimize image contrast. This wide dynamic range of digital radiography with decoupling of the object's density and the optical density of the film might lead to an "unrealistic appearance" of vertebrae on digital X-rays. Conventional radiographs, on the other hand, cannot be

Fig. 2 Digital radiograph with a T-score of −1.69, rated true-positively as osteopenic by all four readers
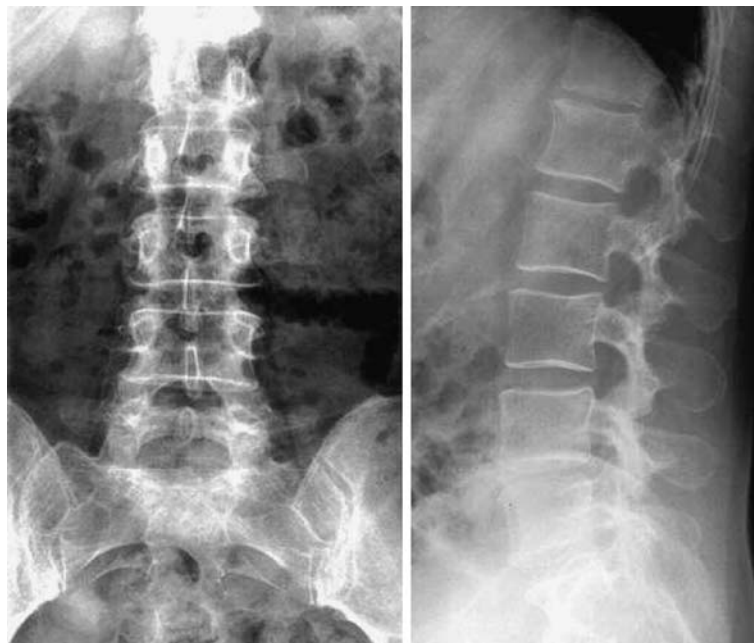
post-processed so that a specific bone mineral density leads to a specific density image on the radiograph.

We found that with increasing bone mineral density loss measured by DXA, an increasing number of patients was correctly diagnosed as true positive with both methods. As expected it appears easier to diagnose images more correctly as the bone density decreases, because radiological signs of osteoporosis become more obvious with progression of the disease. Our finding of a gradual sensitivity increase with decreasing T-score is also confirmed by the results of Jergas et al., who found a higher agreement of readers for low bone mineral densities [15]. False-positive evaluations occurred with all T-values, even with T-values lower than −3 SD. Eighteen percent of digital radiographs with a T-score of less than −3 SD were incorrectly read as normal, while 6% of conventional images with a T-score below 3 SD were read as false negatives. This may be explained by the fact that, in contrast to other studies [16, 17, 15], we excluded patients with vertebral fractures. Therefore, the readers had to orientate their diagnostic process only on the criteria of increased radiotranslucency of bone, pronounced vertebral endplates and reduction of the spongious network. Additionally, the

Fig. 3 Conventional radiograph with a T-score of +1.83, rated true-negatively as normal by all four readers
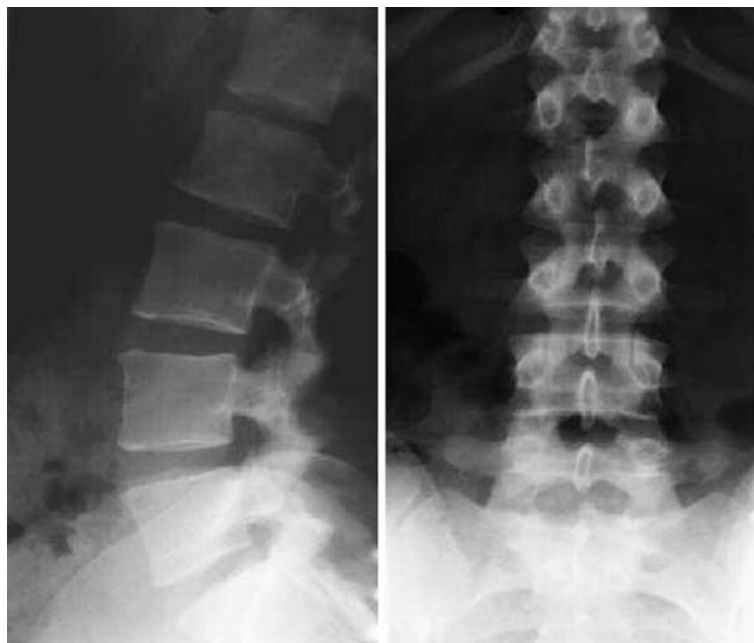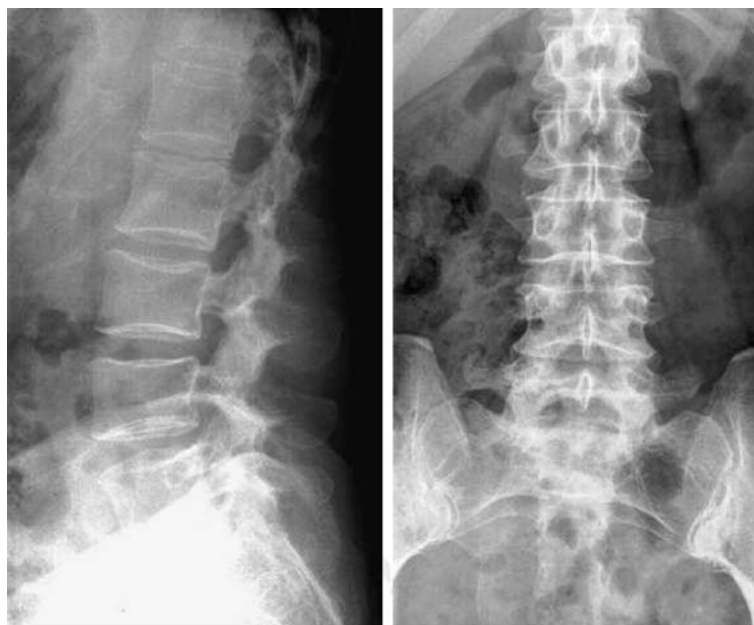
Fig. 4 Conventional radiograph with a T-score of −1.82, rated false-negatively as normal by all four readers

Fig. 5 Digital radiograph with a T-score of −0.86, rated false-positively as osteopenic/osteoporotic by all four readers

readers were blinded to the age and sex of the patients, as this would also bias the diagnostic process. In a study carried out by Ross et al. in 1996, blinding radiographs to film sequence and patient identity led to an increased rate of errors in the detection of vertebral fractures [18].

	Conventional	**Digital**
Sensitivity:	80%	65%
Specificity:	36%	47%

Fig. 6 Results in patients with pathologic bone mineral density (DXA: T-score < -1 SD, osteopenic and osteoporotic patients)

	Conventional	**Digital**
Sensitivity:	96%	82%
Specificity:	36%	47%

Fig. 7 Results in patients with osteopenia (DXA: T-score < -1 SD and > -2.5 SD)

Table 1 Results of the readings of the conventional X-rays compared with DXA (numbers in %)

Conventional images	Specifity	Sensitivity	False negatives	False positives	Ppv	Npv
Reader 1	47.80	89.00	11.00	52.20	75.30	71.00
Reader 2	39.10	80.50	19.50	60.90	70.20	52.90
Reader 3	21.70	84.10	15.90	78.30	65.70	43.50
Reader 4	37.00	89.00	11.00	63.00	71.60	65.40
Mean value	36.40	85.70	14.40	63.60	70.70	58.20

Table 2 Results of the readings of the digital images compared with DXA (numbers in %)

Digital images	Specifity	Sensitivity	False negatives	False positives	Ppv	Npv
Reader 1	65.30	62.40	37.60	34.70	76.50	43.80
Reader 2	49.00	85.30	14.70	51.00	78.80	60.00
Reader 3	28.60	85.30	14.70	71.40	72.60	46.70
Reader 4	44.90	55.00	45.00	55.10	69.00	31.00
Mean value	47.00	72.00	28.00	53.10	74,20	45.40

In this study, for digital as well as for conventional radiography a very low interobserver agreement was found. This might be attributed to the fact that four readers evaluated the images resulting in a lower agreement than an analysis with only two readers. Epstein et al., who carried out a study on the detection of osteoporosis on conventional chest radiographs by two radiologists and one orthopedist, also reported a low interreader agreement [16]. Jergas et al. attributed these low values to the fact that the spine is not easy to evaluate on chest radiograms [15].

However, interobserver agreement, independent if the diagnosis was correct, was much lower for digital radiographs (35%) than for conventional radiographs (73%). Interobserver agreement in combination with a correct diagnosis was also higher on conventional (51%) than on digital (26%) X-rays. This suggests that it is easier to make the correct diagnosis using conventional images, whereas the diagnosis of osteoporosis on digital images is more difficult. One contributing factor might be that a reader more experienced in a particular imaging technique may achieve better results with this technique. Cockshott reported more precise and consistent evaluations by experienced readers than by radiologists with little training [19]. No reports exist in literature concerning whether radiologists trained on conventional radiographs have more difficulties evaluating digital radiographs or vice versa. In our study, two of the readers had more years of radiological experience with conventional than digital radiography, while the other two readers had similar experience with conventional and digital radiographs. However, we found no influence of the readers' experience with a particular technique, and only one reader with more training on conventional radiography had better results with this technique; the other conventionally trained reader had better results on digital radiographs. One of the two readers with equal training with both methods achieved better results on digital, and the other had better results with conventional radiographs.

The low interobserver agreement and even false-negative results in patients with very low T-scores on spinal X-rays indicate that the interpretation of both conventional and digital spinal radiograms in terms of bone mineral density is insufficient in patients without vertebral fractures. Patients without fractures but with osteopenia should be diagnosed before osteoporotic fractures occur, and therefore, the detection rate of osteopenia/osteoporosis on spinal radiographs remains problematic.

Digital and conventional radiography, which are of comparable value in assessing bone mass, but have a low overall success rate, should therefore not be used to direct therapy without first obtaining formal measurement of BMD.

Conclusion

The replacement of conventional radiography by digital X-rays in many radiological departments led to the question whether there are any differences in the detection of osteoporosis on conventional in comparison to digital radiographs. In this study on the detection rate of osteoporosis with digital radiographs compared to conventional radiographs in 286 patients, no significant differences in the diagnostic accuracy of osteopenia/osteoporosis for digital versus conventional radiographs were found. Conventional radiography resulted in a slightly higher sensitivity and a markedly better interobserver agreement than digital radiography, consistent with a more "stable" interpretation of conventional radiographs. In digital radiographs, different image processing and postprocessing algorithms may manipulate the visual aspect of bone density. However, the interpretation of both conventional and digital spinal radiograms in terms of bone mineral density is still insufficient in patients without vertebral fractures, and therefore DXA should be used generously in patients with risk factors for osteoporosis.
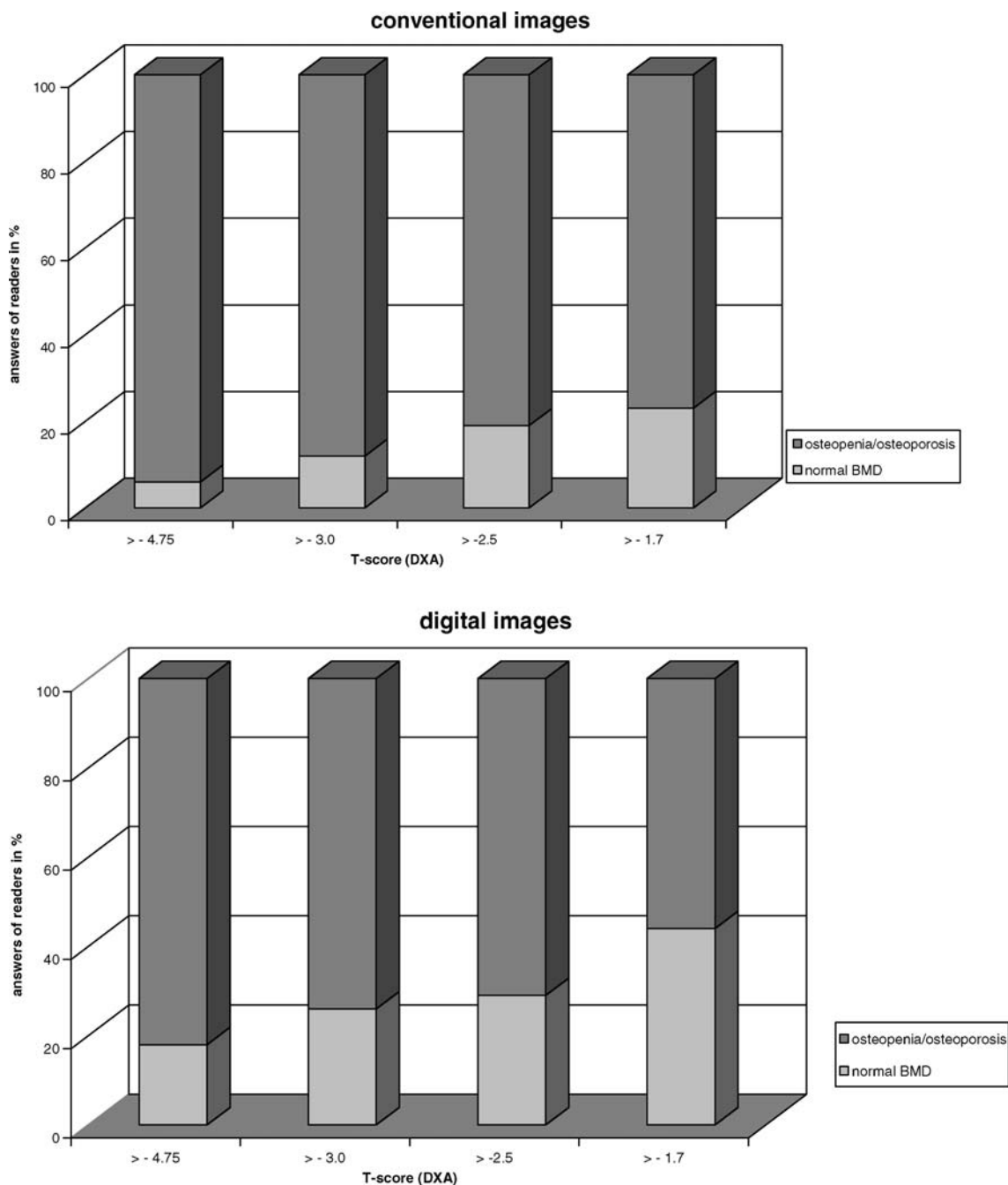
conventional images

digital images

Fig. 8 Gradual increase of sensitivity with falling T-score

References

1. Jergas M, Glüer C, Grampp S, Köster O (1992) Radiologische Diagnostik der Osteoporose. Aktuelle Methoden und Perspektiven. Akt Radiol 2:220–229
2. Keck E (1993) Das Ergebnispapier der "Consensus Development Conference 1993 über Diagnose, Prophylaxe und Behandlung der Osteoporose". Osteologie 2:181–184
3. Jergas M, Genant HK(1993) Current methods and recent advances in the diagnosis of osteoporosis. Arthritis Rheum 36:1649–1662
4. Grigoryan M, Guermazi A, Roemer FW, Delmas PD, Genant HK (2003) Recognizing and reporting osteoporotic vertebral fractures. Eur Spine J 12:104–112
5. Albright F, Smith PH, Richardson AM (1941) Postmenopausal osteoporosis. JAMA 116:2465ff
6. Nathanson L, Lewitan A (1941) Deformities and fractures of the vertebrae as a result of senile and presenile osteoporosis. Am J Roentgenol 46:197–202
7. Stein JA, Lazewatsky JL, Hochberg AM (1987) Dual energy X-ray bone densitometer incorporating an internal reference system. Radiology 165:313

8. Jergas M, Genant HK (1997a) Lateral dual X-ray absorptiometry of the lumbar spine: current status. Bone 20:311–314
9. Cummings SR, Black DM, Nevitt MC, Browner W, Cauley J, Ensrud K, Genant HK, Hulley SB, Palermo L, Scott J, Vogt TM (1993) Bone density at various sites for prediction of hip fractures: the study of osteoporotic fractures. Lancet 341:72–75
10. Jergas M, Genant HK (1997b) Spinal and femoral DXA for the assessment of spinal osteoporosis. Calcif Tissue Int 61:351–357
11. Grampp S, Jergas M, Glüer CC, Lang P, Brastow P, Genant HK (1993) Radiological diagnosis of osteoporosis: current methods and perspectives. Radiol Clin North Am 31:1133–1145
12. World Health Organization (2003) Prevention and management of osteoporosis. WHO Tech Rep Ser 921:1-164
13. Genant HK, Jergas M (2003) Assessment of prevalent and incident vertebral fractures on osteoporosis research. Osteoporosis Int 14:43–55
14. Kovarik J, Küster W, Seidl G, Linkesch W, Dorda W, Willvonseder R, Kotscher E (1981) Clinical relevance of radiological examination of the skeleton and bone density measurements in osteoporosis of old age. Skeletal Radiol 7:37–41
15. Jergas M, Uffmann M, Escher H, Glüer C-C, Young, KC; Grampp S, Köster O, Genant HK (1994) Interobserver variation in the detection of osteopenia by radiography and comparison with dual X-ray absorptiometry of the lumbar spine. Skeletal Radiol 23:195–199
16. Epstein DM, Dalinka MK, Kaplan FS, Aronchick JM, Marinelli DL, Kundel HL (1986) Observer variation in the detection of osteopenia. Skeletal Radiol 15:347–349
17. Williamson MR, Boyd CM, Williamson SL (1990) Osteoporosis: diagnosis by plain chest film versus dual photon bone densitometry. Skeletal Radiol 19:27–30
18. Ross PD, Huang C, Karpf D, Lydick E, Coel E, Hirsch L, Wasnich R (1996) Blinded reading of radiographs increases the frequence of errors in vertebral fracture detection. J Bone Miner Res 11:1793–800
19. Cockshott WP, Park WM (1983) Observer variation in skeletal radiology. Skeletal Radiol 10:86–90