Eur Radiol (2015) 25:2074–2079
DOI 10.1007/s00330-014-3584-0

COMPUTED TOMOGRAPHY

Opportunistic screening for osteoporosis on routine computed tomography? An external validation study

Constantinus F. Buckens · Gawein Dijkhuis · Bart de Keizer · Harald J. Verhaar · Pim A. de Jong

Received: 10 June 2014 / Revised: 23 October 2014 / Accepted: 22 December 2014 / Published online: 17 January 2015

Abstract
Objectives Opportunistic screening for osteoporosis using computed tomography (CT) examinations that happen to visualise the spine can be used to identify patients with osteoporosis. We sought to verify the diagnostic performance of vertebral Hounsfield unit (HU) measurements on routine CT examinations for diagnosing osteoporosis in a separate, external population.
Methods Consecutive patients who underwent a CT examination of the chest or abdomen and had also received a dual-energy X-ray absorptiometry (DXA) test were retrospectively included. CTs were evaluated for vertebral fractures and vertebral attenuation (density) values were measured. Diagnostic performance measures and the area under the receiver operator characteristics curve (AUC) for diagnosing osteoporosis were calculated.
Results Three hundred and two patients with a mean age of 57.9 years were included, of which 82 (27 %) had osteoporosis according to DXA and 65 (22 %) had vertebral fractures. The diagnostic performance for vertebral HU measurements was modest, with a maximal AUC of 0.74 (0.68 – 0.80). At that optimal threshold the sensitivity was 62 % (51 – 72 %) and the specificity was 79 % (74 – 84 %).
Conclusions We confirmed that simple trabecular vertebral density measurements on routine CT contain diagnostic information related to bone mineral density as measured by DXA, albeit with substantially lower diagnostic accuracy than previously reported.
Key Points
- *We externally validated the value of vertebral trabecular bone attenuation for osteoporosis*
- *These diagnostic performance measures were, however, substantially lower than previously reported*
- *This information might be useful when considering the implementation of opportunistic osteoporosis screening*

Keywords Osteoporosis · Vertebral density · Computed tomography · Vertebral fracture · External validation

Electronic supplementary material The online version of this article (doi:10.1007/s00330-014-3584-0) contains supplementary material, which is available to authorized users.

C. F. Buckens · G. Dijkhuis · P. A. de Jong
Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands

B. de Keizer
Department of Nuclear Medicine, University Medical Center Utrecht, Utrecht, The Netherlands

H. J. Verhaar
Department of Geriatric Medicine, University Medical Center Utrecht, Utrecht, The Netherlands

C. F. Buckens (✉)
Department of Radiology, Universitair Medisch Centrum Utrecht, Huispostnummer E01.132, 3508 GA Utrecht, The Netherlands
e-mail: cfmbuckens@gmail.com

Introduction

It has recently been suggested that opportunistic screening for osteoporosis using routine computed tomography (CT) examinations that happen to visualise the spine can be used to identify patients with osteoporosis [1]. Osteoporotic fractures are a major contributor to late life morbidity and mortality and impose a substantial societal cost. Despite the availability of treatments of proven efficacy, there is room for improving the treatment rates at the time of and after major fragility fractures [2–5].

Dual-energy X-ray absorptiometry (DXA) is a widely used tool in assessing osteoporosis. The widespread employment of CT imaging in the course of routine care can be used for opportunistic screening of populations for osteoporosis

outside of any existing screening programs. Risk assessment tools that incorporate multiple clinical parameters such as the World Health Organization's FRAX tool are increasingly commonly used to identify patients who may be at an increased fracture risk. Most CT examinations include all or part of the spine, providing the opportunity to measure vertebral attenuation values of the trabecular regions of vertebral bodies expressed by Hounsfield units (HU) potentially providing a free source of information reflecting bone mineral density (BMD) in a distinct population compared to that are currently being considered for DXA, but only partially overlapping with that population.

The vertebral attenuation values of the trabecular (non-cortical) regions of vertebral bodies, as expressed by Hounsfield units (HU), can be extracted from CT [1]. Trabecular bone is preferentially affected by osteoporosis, particularly in the early phases of the disease process [6]. Additionally, vertebral compression fractures can also be visualized on CT [7, 8]. Following data from pragmatic prognostic studies, these now contribute directly to treatment decisions in new guidelines and form part of the indication to begin treatment, along with clinical history, DXA T-score, and laboratory evaluations [9, 10]. This has led to the inclusion of a lateral view to standard DXA assessment so that vertebral fracture assessment is also possible on DXA.

A threshold of 110 HU for lumbar vertebra 1 (L1) or 115 HU for thoracic vertebra 12 (Th12) was proposed as a cut-off yielding high specificity [11] for identifying patients at risk of osteoporosis and potentially in need of further screening and treatment.

Before further large outcome studies in clinical or screening settings can be pursued it is crucial to determine the external validity of this approach [12]. The objective of this study was to evaluate the diagnostic performance of vertebral HU measurements on routine CT examinations for diagnosing osteoporosis in an external population. For this we have included a population separate to the initial sample in which vertebral attenuation was investigated [1], but one which remains within routine radiological practice: we included a population of convenience who underwent a CT examination of the chest or abdomen and who had also received a DXA test within 90 days.

Materials and methods

Setting and patients

The University Medical Centre Utrecht Ethical Review Board approved this study and the need for informed consent was waived. Consecutive patients who underwent a CT examination of the chest or abdomen between 2005 and 2012 and who had also received a DXA test within 90 days (before or after the CT) were retrospectively included. CT examinations were acquired in the course of routine care using multidetector CT systems (16-256 detector rows, Philips Medical Systems, Best, the Netherlands). Dual-energy X-ray absorptiometry (DXA, Hologic Discovery A, Hologic Inc, Bedford, MA, USA) was performed on the spine (L2 – L4) and hips in the course of routine care. As per common practice, the cut-off for osteoporosis was set as having a T-score ≤-2.5 at any measured location, either in L2 – L4 and/or a hip. The cut-off for osteopenia was ≤-1.

Measurement of vertebral fractures on CT

Sagittal CT reformats were evaluated for vertebral fractures (height loss ≥25 % compared to an adjacent normal vertebra) according to Genant´s semiquantitative Vertebral Fracture Assessment (VFA) method [13, 14]. The reconstructions were assessed at or around the mid-sagittal point. Observers were free to scroll, adjust the orientation and window. This method on CT has previously been shown to have excellent reliability for the presence of a vertebral fracture [7]. One of two observers, either one board certified radiologist or one senior radiology resident with a special interest in musculoskeletal imaging, performed the measurements. They were blinded to the DXA results.

Measurements of bone density on CT

CT attenuation values were measured by the same two observers on stored axial images within trabecular regions of the bodies of L1 or the nearest visible, unfractured, visually normal vertebra, as previously described (Fig. 1) [1]. If L1 was not visualized, Th12 was measured instead. A single click and drag region of interest was placed in the anterior, upper portion of the body of the vertebra, inside trabecular bone. The region of interest was drawn so that it was as large as possible without intersecting the vertebral cortex and without including dense bone islands, hemangiomas, or traversing vessels. The CT attenuation was measured using Hounsfield units, with lower values representing lower bone mineral density.

Statistical analysis

Sensitivity, specificity, positive and negative predictive value, accuracy, and area under the receiver operator characteristics curve (AUC) was calculated for three different pre-selected thresholds (80, 110, and 160 HU) of mean HU for trabecular vertebral bone density for DXA-defined osteoporosis, as previously described [1]. We also determined the optimal HU threshold in this cohort, defined as the threshold yielding the maximum proportion of correct classifications (i.e. true positives and true negatives) from a univariate logistic regression model, where CT attenuation was set as the independent

variable and DXA-defined osteoporosis the dependant variable. A multivariate logistic regression model was fitted to assess the value of adding age and gender to vertebral density on CT when predicting DXA-defined osteoporosis.

Subsequently, these analyses were repeated after adding the variable presence of one or more vertebral fractures to the attenuation classifier at each threshold. In effect, patients with a vertebral fracture were thus classified as having CT-defined 'osteoporosis' regardless of their HU value. For example, a patient with an HU value at L1 of 150 HU would not be considered to have a low HU at the 80 or 110 thresholds. If that patient had a prevalent vertebral fracture, however, this patient would be reclassified as being at risk regardless of having a bone density above the HU thresholds. This approach seeks to incorporate all the available information in a CT for predicting DXA-defined osteoporosis.

Finally, we examined the relation between vertebral fractures and DXA and vertebral density values on CT, whereby vertebral fractures were set as the dependant variable and DXA T-scores and vertebral density on CT were in turn set as the independent variables in two univariate models.

Logistic regression models were fit to ascertain the optimal performance and threshold of vertebral HU in our sample, and to examine the incremental diagnostic value of fracture status, age, and gender in a multivariate model.

Results

Three hundred and two patients (98 men) with a mean (SD) age of 57.9 (15.2) years were included, of which 82 (27 %) had osteoporosis according to DXA, whilst 132 (44 %) were osteopenic and 82 (27 %) had a normal BMD. The mean interval between DXA and CT was 37.1 (26.3) days. Vertebral fractures were found in 65 (22 %) of the patients, 27 (33 %) amongst those with DXA defined osteoporosis, and in 38 (17 %) patients without osteoporosis (Table 1, Fig. 2). Amongst women with DXA-defined osteoporosis, 19/56 (34 %) had a vertebral fracture compared to 28/148 (19 %) amongst those without osteoporosis. Men showed similar proportions, with 8/26 (31 %) with at least one fracture amongst osteoporotic men and 10/72 (14 %) amongst non-osteoporotic. Vertebral fractures were most frequent amongst patients with lower vertebral density values on CT, with all but 61 (94 %) occurring in those with L1 values below 160 HU (Fig. 2).

The diagnostic performance for vertebral HU measurements was modest, as measured by the AUC (Table 2, Fig. 3). We found an AUC of 0.74 (0.68 – 0.080), with an optimal diagnostic performance (the threshold of HU where the proportion of correctly classified patients was the greatest) at the HU threshold of L1 ≤ 99/Th12 ≤ 104. At that threshold, the sensitivity was 62 % (51 – 72) and the specificity was 79 % (74 – 84). At the threshold defined by Pickhardt as achieving the best balance between sensitivity and specificity ($HU \leq 160/HU_{Th12} \leq 165$), we found a high sensitivity of 91 % (84 – 98) with a specificity of 29 % (23 – 35).

Including prevalent vertebral fractures as a dichotomous classifier did not improve the diagnostic performance (Table 2). The optimal AUC was virtually unchanged at 0.74 (0.676 – 0.804), and was slightly lower than the CT vertebral density only model at the three preselected thresholds,

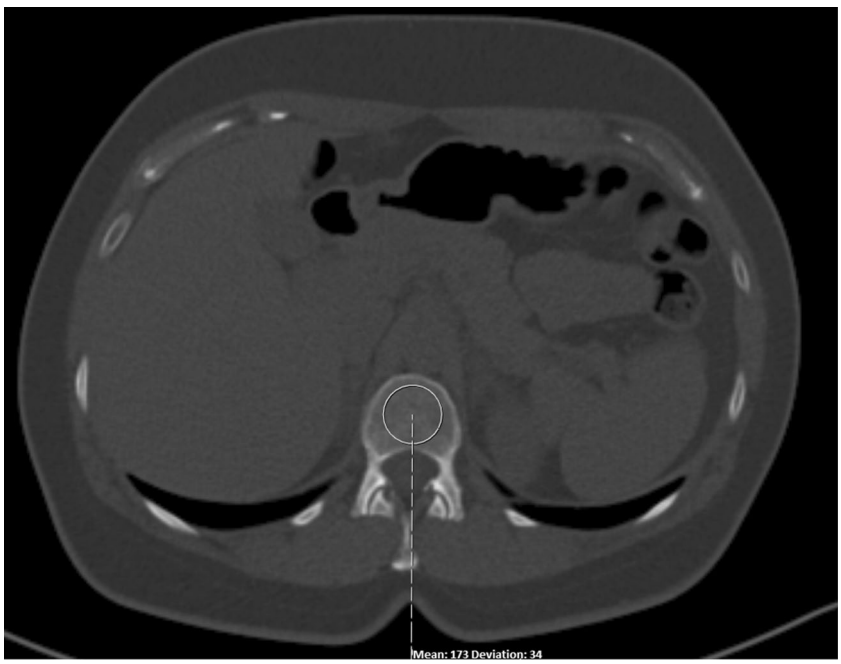

Fig. 1 Axial image through the superior part of the vertebral body of L1 showing placement of a measurement region of interest placed within the trabecular bone

Table 1 Characteristics of patients included with and without osteoporosis, as defined by a DXA T-score <-2.5 *if first lumbar (L1) was not visualized or was fractured, the twelfth thoracic (Th12) was measured instead. **Thorax and abdominal CTs were included. ***lowest areal bone density value obtained value used, either derived from lumbar spine or hip

Variable	Description	Osteoporosis	No osteoporosis
n	Number	82	220
Age	Years (SD)	61 (16)	57 (15)
Sex	Male (%)	26 (32 %)	72 (33 %)
Vertebra measured*	L1 (%)	72 (88 %)	198 (90 %)
CT anatomical area**	Abdomen (%)	53 (65 %)	136 (62 %)
DXA area***	Lumbar spine (%)	58 (71 %)	119 (54 %)
T-score	Mean (SD)	-3.2 (0.62)	-1.1 (0.96)
Days	Mean (SD)	30 (25)	40 (26)
Fracture	Fracture (%)	27 (33 %)	38 (17 %)
HU	Mean (SD)	97 (40)	138 (50)
Indication	Inflammatory/autoimmune	10 (12 %)	41 (19 %)
	Endocrine disorder	12 (15 %)	63 (29 %)
	Fracture	28 (34 %)	44 (20 %)
	Glucocorticoid therapy	9 (11 %)	25 (11 %)
	Transplantation	10 (12 %)	17 (8 %)
	Malignancy	3 (4 %)	7 (3 %)
	Other	10 (12 %)	23 (10 %)
$HU_{L1}80/HU_{Th12}85$		33 (40 %)	21 (10 %)
$HU_{L1}110/HU_{Th12}115$		56 (68 %)	74 (34 %)
$HU_{L1}160/HU_{Th12}165$		75 (91 %)	156 (71 %)

reflecting the lower specificity associated with this approach. At the lowest HU threshold (80/85 HU) the addition of fracture presence reclassified a total of 38 patients, of which 13 had osteoporosis. For the middle and the highest thresholds, fracture presence reclassified 14 (two with osteoporosis) and five of the participants (none with osteoporosis), respectively. This represents a negative net reclassification index at all three thresholds due to lower specificity (Table 2). In a multivariate logistic regression model age (p=0.23), gender (p=0.66), and fracture status (p=0.63) did not significantly improve the model. The AUC of this model was also 0.74 (0.68 – 0.80).

With vertebral fractures as an outcome and DXA T-score as the independent variable, we found an AUC of 0.611 (0.53 – 0.69). When vertebral density on CT was set as the independent variable we found a higher AUC of 0.765 (0.70 – 0.83).

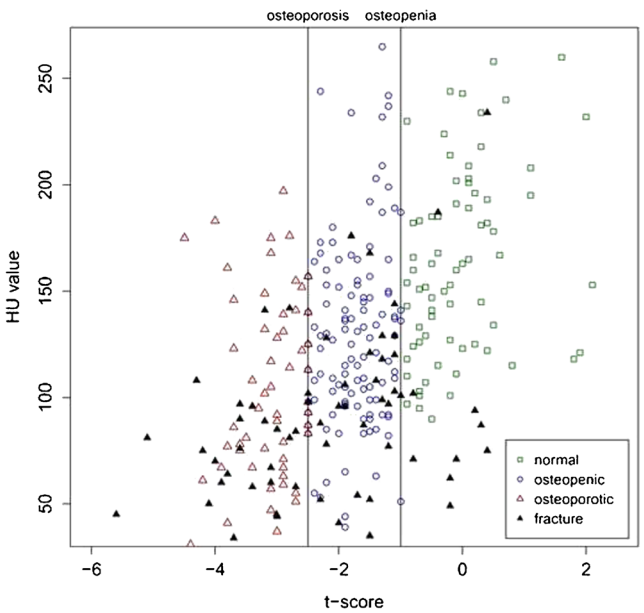

Fig. 2 scatterplot of the vertebral densities in Houndsfield units (HU) and DXA T-scores of the sample population. Normal (*squares*), osteopenic (*circles*), and osteoporotic (*hollow triangles*) individuals are displayed, along with those who had a vertebral fracture (*solid triangles*) on CT

Discussion

We confirmed that simple trabecular vertebral density measurements on routine CT contain diagnostic information related to bone mineral density as measured by DXA, albeit with substantially lower diagnostic accuracy than previously reported.

Pickhardt et al. [1] had found a higher diagnostic performance, with an AUC around 0.83 and a sensitivity of 76 % paired with specificity of 75 % at a 135 HU threshold. These performance measures were substantially higher than the AUC of 0.74 and optimal sensitivity of 62 % and specificity of 79 % we observed (at an optimal L1 ≤ 99/Th12 ≤ 104 threshold). This discrepancy might suggest that the diagnostic performance of vertebral density on CT for DXA-defined

Table 2 Diagnostic accuracy of routine computed tomography for osteoporosis defined as DXA T-score <-2.5. *AUC*, area under the receiver operator characteristics curve

Determinants	Threshold	Sensitivity	Specificity	Accuracy	Positive predictive value	Negative predictive value	AUC
Vertebral density	$HU_{L1}80/HU_{Th12}85$	40 % (30–51)	90 % (86 – 94)	77 % (71-82)	60 % (44–76)	80 % (77–84)	0.65 (0.59-0.71)
	$HU_{L1}110/HU_{Th12}115$	68 % (59-78)	66 % (60-73)	67 % (60-74)	43 % (35-52)	85 % (80-90)	0.67 (0.61-0.73)
	$HU_{L1}160/HU_{Th12}165$	91 % (84-98)	29 % (23-35)	46 % (39-52)	32 % (29-36)	90 % (80-98)	0.60 (0.56-0.65)
	Optimal* $HU_{L1}99/HU_{Th12}104$	62 % (51-72)	79 % (73- 84)	74 % (67-81)	52 % (41-62)	85 % (80-89)	0.74 (0.68-0.80)
Vertebral density and vertebral fracture	$HU_{L1}80/HU_{Th12}85$ and/or fracture	56 % (45-67)	79 % (74-84)	73 % (66-79)	50 % (39-61)	83 % (78-87)	0.68 (0.62-0.74)
	$HU_{L1}110/HU_{Th12}115$ and/or fracture	71 % (61-80)	61 % (55-67)	64 % (57-71)	40 % (33-47)	85 % (79-90)	0.66 (0.60-0.72)
	$HU_{L1}160/HU_{Th12}165$ and/or fracture	91 % (85-96)	27 % (21-33)	44 % (38-50)	32 % (28-35)	89 % (79-96)	0.59 (0.55-0.64)

Data are given with 95 % confidence interval between brackets. Confidence intervals generated using 2,000 stratified bootstrap replicates. *optimal threshold is the threshold where the largest proportion of patients is correctly classified.

osteoporosis is relatively variable across different populations and different settings. Differences in scanning equipment from different manufacturers and differences in scanning protocols might explain a part of the difference, along with differences in the population demographics under study. Although the age (mean overall age 58 years) and gender (68 % female overall) distributions in our sample did not differ substantially from those of Pickhardt (59 years and 81 %, respectively), more subtle differences in referral patterns to DXA and CT are likely to exist. Although this limited study cannot fully explain the differences in diagnostic accuracy, our data suggest some caution is warranted when considering the performance of CT for the opportunistic screening of osteoporosis. Recent work has shown that bone density measurements on CT at other sites, such as the hip [15], are also highly correlated to DXA measurements. Further research on technical of clinical determinants that influence CT bone density measurements seems warranted.

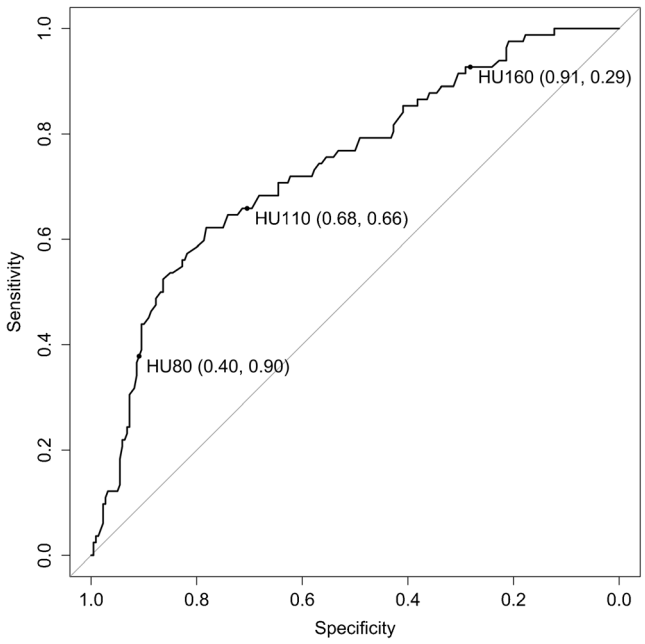

Fig. 3 Receiver operating characteristics curve for vertebral bone density model predicting DXA-assigned osteoporosis. Three predefined thresholds are marked along with the corresponding sensitivity and specificity

We were able to confirm Pickhardt's finding that bone density values as measured from CT have a higher correlation with vertebral fractures than DXA T-score [1]. This suggests that CT density values may be able to complement DXA in assessing deteriorating bone quality, at least in the case of vertebral fractures. The 3D nature of CT data does not suffer from the potential misclassification problem in degenerative and scoliotic spines that periodically occurs when employing DXA, particularly with regard to prevalent vertebral fractures, which may be difficult to detect on DXA in deformed spines is lateral views are not obtained. This is germane as prevalent fractures are important determinants of future facture risk and are an indication to initiate therapy [9, 10]. Longitudinal outcome studies comparing CT to DXA are required to verify this hypothesis, although the primary mode of employing osteoporosis markers on CT is likely to remain in the setting of opportunistically identifying high risk patients in the course of routine diagnostic imaging for the foreseeable future. We observed no effect of age and gender on DXA-determined osteoporosis. This outcome may be a consequence of the source population under study; patients referred for DXA and CT may not meaningfully vary in these variables.

The limitations of this study are the selection biases inherent to the retrospective inclusion of a convenience sample of patients who happen to undergo DXA and CT for a clinical indication. As these patients have a clinical indication to undergo diagnostic testing for osteoporosis, they are likely to represent the population of greatest interest in this field.

Attendant to this limitation, we have not sought any additional clinical information about the included patients beyond that which was included in the imaging referral form. As such nuanced analyses examining, for instance menopausal status and vertebral density on CT, fall beyond the scope of the limited assessment carried out here. Secondly, the cross-sectional nature of this study limits the side-by-side comparison of CT density and DXA for clinical outcome (fracture) and fracture prevention. Although the better correlation of CT density with fractures here suggests that CT density may be better able to capture the deterioration in bone quality that underlies osteoporosis, the prognostic potential of (opportunistic) osteoporosis screening using CT remains to be investigated.

In conclusion, we found that trabecular vertebral density on routine clinical CT correlated well with DXA-values, but found a lower diagnostic performance for DXA-defined osteoporosis than was previously reported. Prospective studies to determine the predictive value of routine CT for future fractures are warranted.

Acknowledgments The scientific guarantor of this publication is Pim De Jong. The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article. This study has received funding by the Netherlands Organization for Scientific Research-Medical Sciences (NWO-MW) grant 40-00812-98-07-005. One of the authors has significant statistical expertise. Institutional review board approval was not required due to previous blanket approval of anonymous retrospective studies on retained diagnostic information that does not impact patient care.

Written informed consent was waived by the institutional review board. Methodology: retrospective, observational, performed at one institution.

References

1. Pickhardt PJ, Pooler BD, Lauder T et al (2013) Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med 158:588–595
2. Bessette L, Ste-Marie L-G, Jean S et al (2008) The care gap in diagnosis and treatment of women with a fragility fracture. Osteoporos Int J Establ Result Coop Eur Found Osteoporos Natl Osteoporos Found USA 19:79–86
3. Metge CJ, Leslie WD, Manness L-J et al (2008) Postfracture care for older women: gaps between optimal care and actual care. Can Fam Physician Méd Fam Can 54:1270–1276
4. King AB, Fiorentino DM (2011) Medicare payment cuts for osteoporosis testing reduced use despite tests' benefit in reducing fractures. Health Aff Proj Hope 30:2362–2370
5. Leslie WD, Giangregorio LM, Yogendran M et al (2012) A population-based analysis of the post-fracture care gap 1996-2008: the situation is not improving. Osteoporos Int J Establ Result Coop Eur Found Osteoporos Natl Osteoporos Found USA 23: 1623–1629
6. Riggs BL, Melton LJ, Robb RA et al (2007) A population-based assessment of rates of bone loss at multiple skeletal sites: evidence for substantial trabecular bone loss in young adult women and men. J Bone Miner Res 23:205–214
7. Buckens CF, de Jong PA, Mol C et al (2013) Intra and interobserver reliability and agreement of semiquantitative vertebral fracture assessment on chest computed tomography. PLoS ONE 8:e71204
8. Buckens C, de Jong P, Mali W et al (2013) Prevalent vertebral fractures on chest CT: higher risk for future hip fracture. J Bone Miner Res Off J Am Soc Bone Miner Res. doi:10.1002/jbmr.2028
9. Raisz LG (2005) Clinical practice. Screening for osteoporosis. N Engl J Med 353:164–171
10. The National Osteoporosis Foundation (NOF) (2013) Clinician's guide to prevention and treatment of osteoporosis. The National Osteoporosis Foundation, Washington, DC
11. Majumdar SR, Leslie WD (2013) Conventional computed tomography imaging and bone mineral density: opportunistic screening or "incidentaloporosis"? Ann Intern Med 158:630–631
12. Altman DG, Royston P (2000) What do we mean by validating a prognostic model? Stat Med 19:453–473
13. Genant HK, Wu CY, van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res Off J Am Soc Bone Miner Res 8:1137–1148
14. Oei L, Rivadeneira F, Ly F et al (2013) Review of radiological scoring methods of osteoporotic vertebral fractures for clinical and research settings. Eur Radiol 23:476–486
15. Gruber M, Bauer JS, Dobritz M et al (2013) Bone mineral density measurements of the proximal femur from routine contrast-enhanced MDCT data sets correlate with dual-energy X-ray absorptiometry. Eur Radiol 23:505–512