Netherlands
The Journal of Medicine

ORIGINAL ARTICLE

Quantitative ultrasound of the heel as triage test to measure bone mineral density compared with dual energy X-ray absorptiometry in men with prostate cancer commencing with androgen deprivation therapy

C. van Casteren-Messidoro[1]*, A.M. Huisman[1], E. Birnie[2,3], M. van Gelder[4], F.E. van de Geijn[1], P. Hamberg[4,5]

Departments of [1]Rheumatology [2]Statistics, [4]Internal Medicine, [5]Prostate Cancer Center, Sint Franciscus Gasthuis, the Netherlands, [3]Institute of Health Policy and Management, Erasmus University, Rotterdam, the Netherlands, *corresponding author: tel.: +31(0)10-4616150, email: chiaramessidoro@hotmail.com

ABSTRACT

Objective: Androgen deprivation therapy (ADT) puts patients at an increased risk of developing osteoporosis. Assessment of bone mineral density (BMD) is most commonly performed by dual energy X-ray absorptiometry (DXA). Alternative ways of estimating BMD, such as quantitative ultrasound (QUS) measurement of the heel, are explored as DXA is expensive, non-portable and uses ionising radiation. We therefore investigated the diagnostic value of QUS as compared with DXA in patients commencing ADT.

Methods: In this cross-sectional study of 60 patients with prostate cancer who were about to start ADT, BMD was measured with DXA and QUS. The fracture risk score, as implemented by the Dutch National Osteoporosis Guideline, was also measured.

Results: No significant correlations were found between the separate DXA T scores and worst DXA T score, and the QUS T scores. Correlations between DXA T scores/QUS scores and fracture risk score were also non-significant. If QUS had been used as a screening tool, with a threshold of T ≤ -0.5 to perform DXA, then relevant osteopenia/osteoporosis (worst DXA T score ≤ -2.0) would have been missed in 1/18 (5.6%) patients. The negative predictive value is 0.95. Using QUS as a screening test prior to DXA and a QUS threshold T score ≤ -0.5 would avoid 21 (35%) DXA scans at the cost of missing one (5.6%) case.

Conclusion: QUS testing cannot replace DXA scans fully as a diagnostic test. However, QUS can be incorporated as triage test prior to DXA to reduce the need for unnecessary DXA scans and the associated costs.

KEYWORDS

Androgen deprivation therapy, osteoporosis, bone mineral density, dual energy x-ray absorptiometry, quantitative ultrasound

INTRODUCTION

Androgen deprivation therapy (ADT), frequently used as part of the treatment of prostate cancer, puts patients at an increased risk for developing osteoporosis, resulting from therapy-induced hypogonadism. Androgen deprivation can be achieved by eliminating the secretion of testicular androgens by surgical castration or by reducing circulating androgens by inhibiting the hypothalamic-hypophyseal-gonadal axis. Hormonal therapy using luteinising hormone releasing hormone (LHRH) agonists or antagonists is currently the main way to achieve medical castration.[1] Hormonal therapy is currently part of the curative treatment of prostate cancer in patients with high-risk prostate cancer who receive a combination of radiotherapy and hormonal therapy, the so-called Bolla schedule;[2] hormonal therapy is also used as palliative treatment in metastatic disease.

During long-term therapy, ADT reduces bone mineral density (BMD) and increases the risk of clinical fractures. During initial ADT, BMD on the hip and spine decrease by approximately 3% per year. Most studies have reported that BMD continues to decline steadily during long-term therapy.[3]

Assessment of BMD is considered the standard evaluation of patients commencing ADT, according to the international guidelines. BMD should be measured every two years if the initial T score is < 1.0, or every year if the T score is between 1.0-2.5, in the absence of associated risk factors. Otherwise, active protective bone treatment should have started at the initiation of ADT.[1,4]

The World Health Organisation defined osteoporosis based upon dual energy X-ray absorptiometry (DXA) measurements, as this is the most widely used method for measuring BMD. DXA provides accurate measurements at clinically relevant sites, i.e. those with major clinical consequences when a fracture occurs. The major disadvantages of DXA are that the instrument is large (not portable), relatively expensive compared with alternative peripheral technologies, and uses radiation albeit at a low dose. Alternative techniques to evaluate bone status at peripheral sites have been developed, e.g. quantitative ultrasound (QUS) measurement of the heel, considered one of the best alternative methods currently available for the assessment of fracture risk.[5]

Replacing DXA measurement by QUS measurements seems attractive as QUS is inexpensive, transportable and free of ionising radiation.[6] Full replacement of a DXA-based diagnostic strategy by QUS measurements is only desirable when the validity of QUS in patients treated with ADT has been demonstrated. In this paper we investigate whether BMD can be validly assessed at the onset of ADT by QUS compared with DXA as reference test. Moreover, if QUS as stand-alone test cannot fully replace DXA, we explore the possibilities of a test strategy that is partially based on QUS. Also, we investigate the value of the fracture risk score compared with DXA and QUS.

PATIENTS AND METHODS

Patients and design

Included in this cross-sectional study were patients who started ADT as treatment for prostate cancer in our clinic between March 2011 and March 2012. There were no exclusion criteria. BMD was measured by DXA scan as well as by QUS, in random order, preferably before and otherwise after initiation of ADT.

BMD measurement by DXA scan was performed in terms of three T scores: at the spine (level L1-L4) and both hips (femoral necks or radius/ulna if hip measurement was impossible due to hip replacement), using a Hologic Discovery DXA scan. We added the worst T score of each DXA measurement. QUS was performed using an Achilles Bone Ultrasonogram (GE Healthcare) which measures the bone stiffness index and calculates a T score. We also recorded the fracture risk score, a tool based on clinical parameters identifying patients at risk for osteoporosis, as implemented by the Dutch National Osteoporosis Guideline (*table 1*).[7] This risk score ranges between 0-12; a higher score indicates higher risk. A DXA scan is advised when the risk score is ≥ 4 points.

Table 1. *Patient's characteristics and distribution of fracture risk score (n = 60)*

Characteristic	
Setting	
Bolla	44 (73.3%)
Metastatic prostate cancer	16 (26.7%)
Type of androgen deprivation	
Gosereline	50 (83.3%)
Degarelix	4 (6.7%)
Busereline	4 (6.7%)
Other	2 (3.4%)
Total fracture risk score (see under)	
0	1 (1.7%)
1	20 (33.3%)
2	25 (41.7%)
3	10 (16.7%)
4+	4 (6.7%)
Median (IQR / range)	2 (1-2/ 0-7)
Weight < 60 kg and/or BMI <20 kg/m² (1 point)	0
Age > 60 years (1 point)	58 (96.6%)
Previous fracture (1 point)	5 (8.4%)
Hip fracture in a parent (1 point)	9 (15%)
Reduced mobility (1 point)	8 (13.4%)
Rheumatoid arthritis (1 point)	0
Secondary osteoporosis (1 point)	9 (15%)
Corticosteroids use (4 points)	2 (3.4%)

Items on fracture risk score:

Weight < 60 kg and/or BMI < 20 kg/m²	1 point
Age > 60 years	1 point
Age > 70 years	2 points
Previous fracture after the age of 50 years	1 point
Hip fracture in a parent	1 point
Reduced mobility	1 point
Rheumatoid arthritis	1 point
Disease or condition associated with secondary osteoporosis (see under)	1 point
Corticosteroids use (> 3 months; 7.5 mg/day)	4 points

Disease or condition associated with secondary osteoporosis:

- Inflammatory bowel disease: Crohn's disease and ulcerative colitis
- Malabsorption
- Chronic inflammatory disorders such spondylarthropathy (ankylosing spondylitis), SLE, sarcoidosis
- Organ transplantation
- Diabetes mellitus
- Untreated hyperthyroidism
- Use of anticonvulsants
- Untreated hyperparathyroidism
- COPD
- Pernicious anaemia
- Low sun exposure

Diagnostic strategies

We compared the validity of the following diagnostic strategies: 1) QUS as stand-alone test was compared with DXA T scores considered as reference test; 2) QUS as screening test with the presence of severe osteopenia as reference (defined as DXA T score ≤ -2.0); 3) fracture risk score as screening test with the presence of severe osteopenia as reference (defined as DXA worst T score ≤ -2.0); and 4) a diagnostic index based on the weighted QUS T score and the fracture risk with the presence of osteoporosis as reference.

Statistical analysis

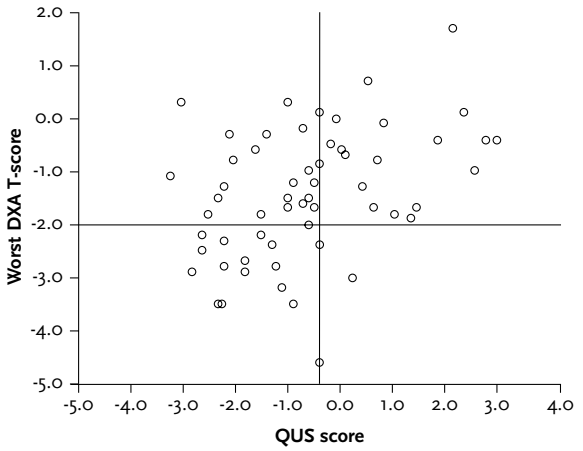
Patient, disease and treatment characteristics at baseline, the QUS and DXA T scores and the differences in T scores between QUS and DXA were described with conventional descriptive statistics: n (%) for nominal and ordinal variables, mean (SD) for quantitative variables with approximate normal distributions and median (interquartile range [IQR]) for quantitative variables with skewed distributions. Intraclass correlation coefficients (ICCs, two way mixed, single measures) were used to quantify the correlations between the QUS T score and the single T scores and worst T score. The strength of associations between the DXA and QUS T scores on the one hand and the fracture risk score was estimated with linear regression analyses and expressed as adjusted R2. The diagnostic strategies were evaluated compared with their reference in terms of the proportion of correct predictions, sensitivity, specificity, positive and negative predictive values (all with 95% CIs). The value of the fracture risk score in addition to the QUS T score was estimated with binary logistic regression and expressed as the change in diagnostic accuracy and the change in the -2log likelihood (-2LL) goodness of fit measure. A p value (two-tailed) < 0.05 was considered a statistically significant difference.

Table 2. *Description of quantitative ultrasound and dual energy X-ray absorptiometry values*

	Mean (SD)	Median	Range Min / Max	IQR
QUS	-0.72 (1.50)	-0.75	-3.3 / 2.8	-1.90 / 0.08
DXA				
T score right hip	-0.40 (0.91)	-0.45	-2.8 / 1.7	-0.90 / 0.30
T score left hip	-0.33 (0.94)	-0.35	-3.1 / 1.7	-0.98 / 0.18
T score LVB	-0.49 (1.59)	-0.70	-3.5 / 3.9	-1.60 / 0.50
T score worst	-1.42 (1.21)	-1.40	-4.6 / 1.7	-2.28 / -0.53

LVB = lumbar vertebral body.

Figure 1. *Scatter plot (worst DXA T score and QUS score)*



RESULTS

The characteristics of the 60 included patients are shown in *table 1*.

Comparison of T scores

The mean QUS value was -0.72. The mean DXA T scores for right hip, left hip and lumbar spine were -0.40, -0.33 and -0.49, respectively. The mean value of the worst DXA T score was -1.42 (*table 2*). The ICCs between all the separate T scores as well as the worst T score with the QUS-derived score were low and not significant. The QUS T score is neither a precise nor a valid estimate of the worst DXA T score (*figure 1*).

Comparison of test characteristics

If QUS had been used as screening tool, with a threshold of T ≤ -0.5 to perform a DXA, then relevant osteopenia (worst DXA T score ≤ -2.0) would have been missed in 1/18 (5.6%, 95% CI 0.1-27.3%) patients, with a negative predictive value (NPV) of 0.95 (95% CI 75.1-99.9%, *table 3*). At a threshold of QUS of T ≤ -0.7 the NPV would be 0.89 (95% CI 70.8-97.7%). Using QUS as triage test prior to DXA and a QUS threshold T score ≤ -0.5, 35% (95% CI 23.1-48.4%) of the DXA scans would have been avoided (*table 3*).

The majority of patients (58.4%) had a low-intermediate fracture risk score of 1-2 or 3 points (scale: 0-12). Only four patients (6.7%) had a risk score of ≥ 4 points (*table 1*). ICCs between DXA T scores/QUS score and fracture risk score resulted in adjusted R2 between -0.014 and -0.017 (all non-significant). The proportion of correct predictions using the fracture risk score of 4 as a cut-off level was 22/60 (36.7%, 95% CI 24.6-50.1%) with a sensitivity of 0/18 (0%, 95% CI 0.0-18.5%), a specificity of 38/42 (90.5%,

Table 3. *Relevant osteopenia/osteoporosis cases with QUS threshold T= -0.5*

	Osteoporosis + (DXA T score ≤ -2.0)	Osteoporosis – (DXA T score > -2.0)	Total
QUS T score ≤ -0.5	17	22	39
QUS T score > -0.5	1	20	21
Total	18	42	60

Sensitivity: 17/18 (94%, 95% CI 73-99%); specificity 20/42 (47%, 95% CI 32-63%).

Table 4. *Relevant osteopenia/osteoporosis cases with fracture risk score*

	Osteoporosis + (DXA worst T score ≤ -2.0)	Osteoporosis – (DXA worst T score ≥ -2.0)	Total
Fracture risk score 0-3	18	38	56
Fracture risk score ≥4	0	4	4
Total	18	42	60

Positive predictive value = 0/4 (0%, 95% CI 0.0-60.2%).

95% CI 0.774-0.973) and an NPV of 38/56 (67.9%, 95% CI 54.0-79.7%) (*table 4*). Hence, QUS does outperform the fracture risk score as triage test.

A diagnostic index based on the fracture risk score and the QUS T score jointly did not improve the diagnostic accuracy nor the -2LL goodness of fit measure compared with a model without the fracture risk score (proportion of correct predictions: 42/60 for both models; -2LL 61.1 vs. 61.5).

DISCUSSION

In our cohort, QUS appears unable to reflect the BMD as generated by DXA, the reference test for the detection of osteoporosis. One plausible explanation for the low correlation is the fact that both methods structurally evaluate different parameters of BMD. DXA measures bone mineral content and bone area and estimates areal BMD. BMD measured by QUS, in contrast, provides no actual measurement of BMD but is the result of measurements derived from the transmission of ultrasound through bone. Measurements are most commonly made at the calcaneus, a skeletal site composed primarily of cancellous (trabecular) bone, similar to the spine. However, the correlation between DXA lumbar spine T score and QUS score in our cohort was one of the lowest and therefore does not support that view. A study of 230 postmenopausal women also showed lower correlation coefficients between QUS and DXA T scores of the lumbar spine compared with the hips, which might be caused by using only one composite T score for the lumbar spine instead of four different values (L1, L2, L3 and L4).[8] A review of comparative studies between bone densitometry and QUS of the calcaneus in osteoporosis reported a sensitivity of QUS compared with DXA in detecting osteoporosis of 65-67% when compared with spine BMD and 72-74% when compared with hip BMD.[9]

As the DXA scan delivers T scores from different sites with a higher likelihood of detecting osteoporosis, the outcome of an alternative screening method should be compared with the worst T score from the DXA scan. In our study the correlation between QUS T score and the worst DXA T score appeared to be the lowest of all, confirming the lack of potential of QUS to assess BMD in our cohort, when compared with the gold standard (*figure 1*).

Many studies have demonstrated that low BMD measured by DXA at any skeletal site (spine, hip of forearm) is predictive for an osteoporotic fracture.[10-12] QUS, however, has also been shown to be a good predictor of osteoporotic fracture risk.[13-22] Advantages of QUS include lower expense, portability and lack of radiation exposure. However, a point of controversy with regard to earlier studies is the cut-off point for the diagnostic determination of osteoporosis with the QUS method. No agreement could be found between the threshold accepted for DXA (a T score < -2.5 for osteoporosis) and a QUS threshold for detecting osteoporosis.[9] A meta-analysis of 25 studies that evaluated the sensitivity and specificity of calcaneus ultrasound for identifying patients with DXA T scores < -2.5 concluded that the currently used ultrasound cut-off thresholds do not have sufficiently high sensitivity or specificity to definitively exclude or confirm DXA diagnosed osteoporosis.[23] However, our cohort concerned patients with a low risk of osteoporosis (due to the low fracture risk score). The majority of patients (62%) had their BMD measurement prior to or within the first two months of ADT, and this makes an ADT-induced effect on BMD less likely.

We also estimated the test characteristics of QUS for different QUS T score thresholds to perform a DXA, with the idea of using QUS as a screening tool. With a QUS threshold of T ≤ -0.5 to perform a DXA we found a high NPV of 0.95, meaning that relevant osteopenia/osteoporosis would have rarely been missed. Using QUS as screening or triage test prior to DXA and a QUS threshold T score ≤ -0.5 would avoid 35% of DXA scans at the cost of missing one (5.6%) relevant case of osteopenia/osteoporosis. A previous review of seven studies comparing different QUS thresholds and DXA T scores resulted in QUS sensitivity ranging from 79-93% and specificity from 28-90% (when at a lower threshold). QUS thresholds had a

variability of -1.7 and -2.4, and a T score of < -3.65 for QUS was equivalent to a T score < -2.5 for DXA.[9]

Besides a benefit in patient comfort by using the mobile device as a screening tool, there is also a relevant cost benefit. A true cost analysis is hampered by the lack of a standardised price for a QUS measurement. An estimate of price based upon a unit price of a QUS machine (€ 16,000) with a full depreciation in ten years, yearly service costs and costs for staff (15 minutes per measurement) is € 13 per measurement. A set price for conventional DXA is € 167 per measurement.

If the standard strategy had been used in our 60 patients, then the costs of the BMD measurement for the whole group would have been € 10,020. The screening strategy would have resulted in 60 QUS measurements and 39 DXA measurements, with total costs of € 7293, rendering a cost reduction of 27%.

This fracture risk score aimed to identify the patients at risk for osteoporotic fractures on the basis of clinical parameters appeared to be a poor predictor of worst DXA T score. We found very weak correlations between the risk score and the DXA T scores and QUS scores. Moreover, the fracture risk score did not identify any of the patients with a relevant osteopenia based on worst DXA T score in our cohort. In summary, the risk score seems to be a poor test, worse than the lower threshold QUS. Jointly, the fracture risk score and the QUS T score did not improve the diagnostic accuracy.

This study has a number of limitations, such as the relatively small number of included patients. Also, we correlated the DXA T scores with the QUS T scores, while there is no consensus on the QUS threshold to predict osteoporosis. Thus the choice to test the sensitivity of QUS with different thresholds, compared with DXA-proven severe osteopenia, which showed a very good NPV of QUS used as a screening tool prior to DXA. This is a retrospective cohort, but prospective studies analysing whether the BMD decline of patients as measured by DXA is predictable by QUS are needed. Also the further role of QUS as a possible screening tool needs to be further elucidated.

In conclusion, QUS is not suitable to detect osteoporosis at the start of ADT treatment in men with prostate cancer, compared with the gold standard (DXA), and therefore cannot replace DXA fully. However, QUS can be incorporated in the diagnostic strategy as triage test prior to DXA to reduce the need for DXA scans and performs better than a risk score based on clinical parameters.

DISCLOSURES

The authors declare no conflicts of interests. No funding or financial support was received.

REFERENCES

1. Mottet N, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part II: Treatment of advanced, relapsing, and castration-resistant prostate cancer. Eur Urol. 2011;59:572-83.
2. Bolla M, Van Tienhoven G, Warde P, et al. External irradiation with or without long-term androgen suppression for prostate cancer with high metastatic risk: 10-year results of an EORTC randomised study. Lancet Oncol. 2010;11:1066-73.
3. Isbarn H, Boccon-Gibod L, Carroll PR, et al. Androgen Deprivation Therapy for the Treatment of Prostate Cancer: Consider Both Benefits and Risks. Eur Urol. 2009;55:62-75.
4. Higano CS. Bone loss and the evolving role of bisphosphonate therapy in prostate cancer. Urol Oncol. 2003;21:392-8.
5. Faulkner KG. Screening for osteoporosis with heel ultrasound. GE Healthcare.
6. Marin F, Gonzalez-Macias J, Diez-Perez A, et al. 2006 Relationship between bone quantitative ultrasound and fractures: a meta-analysis. J Bone Miner Res. 21:1126e1135.
7. http://www.diliguide.nl/document/1015
8. Boyanov M, Shinkov A, Nestorova R. Bone density measurement: quantitative ultrasound of the calcaneus and distal radius. A comparison with dual spectrum X-ray absorptiometry. Dtsch Med Wochenschr. 2007.
9. Floter M, Bittar CK, Zabeu JL, Carneiro AC. Review of comparative studies between bone densitometry and quantitative ultrasound of the calcaneus in osteoporosis. Acta Reumatol Port. 2011:36:327-35.
10. Johnell O, Kanis JA, Oden A, et al. Predictive value of BMD for hip and other fractures. J Bone Miner Res. 2005;20:1185.
11. Leslie WD, Tsang JF, Caetano PA, et al. Effectiveness of bone density measurement for predicting osteoporotic fractures in clinical practice. J Clin Endocrinol Metab. 2007;92:77.
12. Bauer DC, Glüer CC, Genant HK, Stone K. Quantitative ultrasound and vertebral fracture in postmenopausal women. Fracture Intervention Trial Research Group. J Bone Miner Res 1995;10:353.
13. Schott AM, Weill-Engerer S, Hans D, et al. Ultrasound discriminates patients with hip fracture equally well as dual energy X-ray absorptiometry and independently of bone mineral density. J Bone Miner Res. 1995;10:243.
14. Khaw KT, Reeve J, Luben R, et al. Prediction of total and hip fracture risk in men and women by quantitative ultrasound of the calcaneus: EPIC-Norfolk prospective population study. Lancet. 2004; 363:197.
15. Bauer DC, Glüer CC, Cauley JA, et al. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. Study of Osteoporotic Fractures Research Group. Arch Intern Med. 1997;157:629.
16. Hans D, Dargent-Molina P, Schott AM, et al. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. Lancet. 1996;348:511.
17. Glüer CC, Eastell R, Reid DM, et al. Association of five quantitative ultrasound devices and bone densitometry with osteoporotic vertebral fractures in a population-based sample: the OPUS Study. J Bone Miner Res. 2004;19:782.
18. Stewart A, Torgerson DJ, Reid DM. Prediction of fractures in perimenopausal women: a comparison of dual energy x ray absorptiometry and broadband ultrasound attenuation. Ann Rheum Dis. 1996; 55:140.
19. Thompson P, Taylor J, Fisher A, Oliver R. Quantitative heel ultrasound in 3180 women between 45 and 75 years of age: compliance, normal ranges and relationship to fracture history. Osteoporos Int. 1998;8:211.
20. Marín F, González-Macías J, Díez-Pérez A, et al. Relationship between bone quantitative ultrasound and fractures: a meta-analysis. J Bone Miner Res. 2006;21:1126.
21. Dobnig H, Piswanger-Sölkner JC, Obermayer-Pietsch B, et al. Hip and non-vertebral fracture prediction in nursing home patients: role of bone ultrasound and bone marker measurements. J Clin Endocrinol Metab. 2007;92:1678.
22. Nayak S, Olkin I, Liu H, et al. Meta-analysis: accuracy of quantitative ultrasound for identifying patients with osteoporosis. Ann Intern Med. 2006;144:832.
23. Díez-Pérez A, Marín F, Vila J, et al. Evaluation of calcaneal quantitative ultrasound in a primary care setting as a screening tool for osteoporosis in postmenopausal women. J Clin Densitom. 2003;6:237.