Estimation of femoral neck bone mineral density by ultrasound scanning: Preliminary results and feasibility

Paola Pisani a, Francesco Conversano a, Fernanda Chiriacò a, Eugenio Quarta b, Laura Quarta b, Maurizio Muratore b, Aimè Lay-Ekuakille c, Sergio Casciaro a,

⇑

a National Research Council, Institute of Clinical Physiology, Lecce, Italy
b O.U. of Rheumatology, Galateo Hospital, San Cesario di Lecce, ASL-LE, Lecce, Italy
c University of Salento, Department of Innovation Engineering, Lecce, Italy

Article info

Article history:
Received 26 March 2015
Received in revised form 5 August 2016
Accepted 11 August 2016
Available online 23 August 2016

Keywords:
Ultrasound
Femoral neck
Osteoporosis
Biomedical signal processing
Biomedical image processing
Biomedical measurements

abstract

Aim of this paper was to assess the diagnostic accuracy of a novel ultrasound (US) approach for femoral neck densitometry. A total of 173 female patients (56–75 years) were recruited and all of them underwent a dual X-ray absorptiometry (DXA) of the proximal femur and an US scan of the same anatomical district. Acquired US data were analysed through a novel algorithm that performed a series of spectral and statistical analyses in order to calculate bone mineral density employing an innovative method. Diagnostic accuracy of US investigations was quantitatively assessed through a direct comparison with DXA results. The average diagnostic agreement resulted pretty good (85.55%), with a maximum (88.00%) in correspondence of the youngest investigated patients (56–60 y). Overall, diagnostic accuracy showed only minimal variations with patient age, indicating that the proposed approach has the potential to be effectively employable for osteoporosis diagnosis in the whole considered age interval.

1. Introduction

Osteoporosis is the most common bone disease in humans, characterized by a low bone mass and a micro-architectural deterioration of bone tissue, with a subsequent increase in bone fragility and susceptibility to fracture, and representing a major public health problem [1,2]. This pathology affects more than 200 million people worldwide, causing over 8 million of new fractures each year; in Europe, almost 3 million of new osteoporotic fractures occur yearly, causing 43,000 deaths and accounting for a direct cost of about €40 billion [3]. The most frequent osteoporotic fractures occur at either spine or proximal femur, with the latter in particular representing a very common injury for elderly patients, requiring expensive therapies and/or surgeries and frequently resulting in reduced quality of life, disability and mortality [4]. The incidence of femoral fractures increases with age, with a 75% occurring in women [5], and typically accounts for more than 70% of total direct costs of osteoporotic fractures [6]. The mortality rates associated with femoral fractures within 1 year vary from 8% to 36%, depending on concomitant risk factors (age, comorbidity, pre-fracture functional status, etc.) [7], with a higher mortality in men than in women [8]. In addition, femoral fractures are followed by a 2.5-fold increased risk of future osteoporotic fractures [9] and only 40% of fractured patients fully regain their pre-fracture level of independence [2,10].

Taking into account the global increase in life expectancy, which is likely to worsen the situation, the only possible way to reduce the occurrence of femoral fractures is represented by the adoption of more effective strategies for early osteoporosis diagnosis and fracture prevention through population mass screenings. In fact, there is a large gap between the numbers of women that are treated compared to the proportion of the population that could be eligible for treatment based on actual fracture risk [11]. It should be definitely raised the awareness that osteoporosis is actually preventable and treatable, but, since there are no warning signs prior to a fracture, many people are not being diagnosed in time to receive effective therapy during the early phase of the disease [6].

Currently, dual X-ray absorptiometry (DXA) of proximal femur and lumbar spine is the state-of-the-art technique to measure bone mineral density (BMD) and to establish an osteoporosis diagnosis according to the World Health Organization (WHO) guidelines [12]. In particular, femoral neck BMD is associated with a high gradient of risk for femoral fracture [13] and the WHO fracture risk assessment tool (FRAX®) employs the femoral neck BMD as a refer-
ence standard value, which is then integrated with clinical risk factors in order to determine the 10-year fracture probabilities [14]. However, DXA cannot be used in primary healthcare neither for screening purposes because of intrinsic limitations, such as radiation-related issues, high costs, large size of the equipment and limited availability [15]. As a consequence, DXA examination is indicated only in women aged 65 years and older, as well as in younger and peri-menopausal women presenting specific risk factors for fragility fractures [16], and in men aged 70 years and older, or even younger but presenting risk factors for fracture [6].

Over the past ten years, in a period that has seen a significant proliferation of ultrasound (US) applications in the biomedical field because of their fundamental advantages over competing technologies [17–30], US methods have been developed also for osteoporosis diagnosis and fracture risk prediction, aiming at the introduction of non-ionizing and cost-effective bone assessments, integrating BMD estimations and evaluations of micro-structural and elastic properties, which have an important direct influence on actual bone strength [31–33]. However, all the commercially-available US devices can be applied only to peripheral bone districts (calcaneus, phalanges, tibial shaft and radius) and their results present poor correlations with femoral neck BMD as measured by DXA [34].

In a recent conference paper [35], we introduced the preliminary clinical validation of a new US-based methodology for bone densitometry that can be applied directly on femoral neck and showed an appreciable correlation with site-matched DXA outcomes. In the present study we assessed the performance of the proposed method on a larger study population belonging to a wider age interval. Diagnostic accuracy as a function of patient age and general clinical usefulness of the new approach are critically discussed taking into account the most recent literature available papers. Full details of the adopted protocol for data acquisition and processing are also provided and commented.

2. Materials and methods

2.1. Patients

The study was conducted at the Operative Unit of Rheumatology of “Galateo” Hospital (San Cesario di Lecce, Lecce, Italy). A total of 173 consecutive female patients were enrolled, according to the following inclusion criteria: Caucasian ethnicity, aged in 56–75 y, body mass index (BMI) < 40 kg/m², absence of significant deambulation impairments, medical prescription for a femoral DXA. All the recruited patients underwent two different diagnostic investigations: a conventional DXA of the proximal femur and an US scan of the same bone district, as detailed in the next paragraphs.

The study protocol was approved by the hospital ethics review board and all patients gave their informed consent.

2.2. DXA measurements

DXA scans were performed on the proximal femur employing a Discovery W scanner (Hologic, Waltham, MA, USA). In addition to the femoral neck BMD value, expressed as grams per square centimetre (g/cm²), DXA report also included the T-score value, defined as the number of standard deviations (SDs) from the peak BMD of young women found in the standard Hologic reference database for Caucasian women, and the Z-score value, defined as the number of SDs from the BMD of age-matched women found in the same standard reference database. According to the commonly used WHO definitions, patients were classified as “osteoporotic” if T-score ≤ −2.5, “osteopenic” if −2.5 < T-score < −1.0 or “healthy” if T-score ≥ −1.0 [12,36].

2.3. US acquisitions

US scans of the proximal femur were performed using an innovative US device developed in Lecce (Italy) within the ECHOLIGHT Project through a collaboration between CNR-IFC (National Research Council – Institute of Clinical Physiology) and Echolight srl. The device was equipped with a 3.5-MHz broadband convex transducer and configured to provide both echographic images and “raw” unfiltered radiofrequency (RF) signals.

Each patient underwent a proximal femur scan that lasted about 40 s and generated 50 frames of RF data, digitized at 40 MS/s (16 bits), which were acquired and stored in a PC hard-disk for subsequent off-line analysis. Transducer focus and scan depth were specifically adjusted for each acquisition in order to have the femoral neck interface located in the US focal region and in the central part of the image. The other acquisition parameters were kept constant to the following values: power = 75%, mechanical index (MI) = 0.4, gain = 0 dB, linear time gain compensation (TGC). A picture taken during an US acquisition is shown in Fig. 1.

2.4. US data analysis

Acquired US data were analysed through a novel automatic algorithm that performed a series of spectral and statistical analyses, involving both the echographic images and the underlying RF signals, in order to calculate a new US parameter, called “osteoporosis score” (O.S.). The calculation of this parameter for lumbar spine acquisitions has been detailed in a very recent paper [37], in which a strong correlation between O.S. and DXA measured lumbar BMD was also found. The present work, for the first time, illustrates the details of O.S. calculation on femoral neck.

The implemented algorithm performs diagnostic calculations on RF signal segments corresponding to a specific region of interest (ROI) internal to the femoral neck region, which is automatically identified by the algorithm in each acquired frame. Each selected RF signal segment consists of a 200-point Hamming-windowed signal portion starting after the echo from the femoral neck surface, when the amplitude of RF signal envelope reached 15% of its peak value.

The aim of such calculations is to measure the percentage of femoral neck segments whose signal spectral features correlate better with those of an osteoporotic bone model rather than with those of a healthy one. The algorithm actually compares RF spectra calculated from the considered patient dataset with reference models of healthy and osteoporotic femoral necks obtained from previous US acquisitions on DXA-classified patients.

The implementation of the adopted algorithm, which is analogous to the one that has been described in a very recent paper
focused on the application of a similar approach to lumbar spine [37], is herein summarized with specific attention to detail those points in which the adopted algorithm for femoral neck application differs from that described in the referred paper, dedicated to vertebral application.

The main data analysis steps performed on each patient dataset are the following:

1. Automatic identification of proximal femur profile and femoral neck interface within the acquired echographic images.
2. For each femoral neck image, automatic identification of a specific RF signal portion for each scan line crossing the bone surface.
3. Classification of each RF signal portion as “osteoporotic” or “healthy” on the basis of the correlation between its frequency spectrum and each of the two age-matched models stored in a previously obtained reference database.
4. For each frame, calculation of the O.S. value, defined as the percentage of the analysed femoral neck segments that were classified as “osteoporotic” in the previous step.
5. Calculation of the O.S. value for the considered patient as the average of the single frame values.
6. Calculation of the conventional parameters BMD, T-score and Z-score, as a function of the O.S. value, through specific equations depending on patient age and BMI.

Patients enrolled for the present study were subdivided into four different age intervals: 56–60 y, 61–65 y, 66–70 y, 71–75 y. For each of these age intervals, a pair of reference spectral models (an “osteoporotic” one and a “healthy” one) was available in a database that had been previously built following the same procedure detailed in [37] and applying it to US data obtained from femoral acquisitions.

For a generic patient dataset, once the appropriate spectral models had been identified in the reference database, the first operation performed by the algorithm was the automatic segmentation of the proximal femur profile in each acquired image. This was achieved by carrying out the following steps on each considered frame:

- Rearrangement of image data in a rectangular matrix, in order to simplify the subsequent processing steps (the typical acquired image was composed of 253 scan lines having from 4500 to 6000 points/line, depending on the scan depth).
- Brightness masking, aimed at increasing the brightness of the central region while gradually attenuating brightness level toward image boundaries (a custom-designed brightness mask was employed to emphasize the central image portion along the vertical direction).
- Contrast enhancement and image smoothing, implemented through the following sequence: after having normalized pixel values in the range between “0” and “1”, a contrast-limited adaptive histogram equalization (the image was divided into 64 rectangular regions called “tiles”, each tile’s histogram was equalized and the neighbouring tiles were then combined using a bilinear interpolation), followed by a two-dimensional low-pass Gaussian filter (size = 100 × 100, SD = 10) and a further contrast-limited adaptive histogram equalization.
- Histogram equalization on the entire image.
- Thresholding, in order to transform the image into a binary map (threshold value = 0.985).
- Morphologic evaluations, aimed at verifying whether among the white pixel clusters present in the thresholded image was there a “possible femoral profile”, which is a cluster of white pixels that has the typical geometrical features of a proximal femur interface in terms of shape, length, thickness and position (the most strict requirement was the presence of the typical “semicircle” corresponding to femoral head, which, on one side had to present an almost linear extension corresponding to femoral neck and trochanter).

- Selection of the femoral neck interface within the identified proximal femur profile: the identified profile was interpolated by a 13th-order polynomial, which presented a characteristic inflection point in correspondence of the boundary between femoral head and femoral neck, the subsequent inflection point was then assumed as representative of the boundary neck/trochanter, and the femoral neck interface was identified as the tract between the two inflection points (Fig. 2 shows a typical echographic image containing the proximal femur profile with the identification of the three main anatomical sub-regions).
- Spectral validation, consisting in a check of the RF data corresponding to the ROI selected below the femoral neck interface identified in the previous step, in order to verify if the associated spectral content resembled the typical features of a bone structure (i.e., if at least 70% of the spectra obtained from the identified ROI had a Pearson correlation coefficient $r > 0.85$ with at least one of the appropriate reference model spectra).

Once the listed steps had been performed on all the frames belonging to the analysed patient dataset, the algorithm proceeded to the following diagnostic calculations on the RF signals corresponding to the ROIs selected under the identified femoral neck interfaces. The frequency spectrum of each RF signal portion belonging to the considered ROI was classified as “osteoporotic” if the value of its Pearson correlation coefficient with the appropriate osteoporotic model ($f_{\text{oste}}$) was higher than the corresponding correlation value with the related healthy model ($f_{\text{heal}}$), otherwise it was classified as “healthy”. Then, the O.S. value for the considered frame $f_i$ was calculated through the following formula:

$$O.S_{f_i} = \frac{E_{f_{\text{oste}}}}{E_i} \times 100$$

where $E_{f_{\text{oste}}}$ = number of spectra classified as “osteoporotic” for the ROI identified in the frame $f_i$.

$E_i$ = total number of spectra belonging to the ROI identified in the frame $f_i$.

The O.S. value for the considered patient $k$ is:

$$O.S_k = \frac{\sum_{i=1}^{n_k} O.S_{f_i}}{n_k}$$

where $n_k$ represents the number of frames acquired on the patient $k$ and containing an identified femoral neck interface.

Finally, the obtained O.S. value was used as an input parameter to calculate the US-estimated values of BMD, T-score and Z-score.

![Fig. 2](image-url)
through mathematical equations incorporated in the reference model database and whose analytical expressions depended on the specific combination of patient age and BMI.

Diagnostic accuracy of the obtained results was evaluated through a direct comparison with the corresponding DXA values. Every patient dataset was independently included in a specific diagnostic category (osteoporotic, osteopenic, or healthy) by each employed diagnostic technique (i.e., DXA and US): patients that received the same classification by both the systems were considered as "correct diagnoses".

A scheme of the approach used for the comparison of US and DXA evaluations is reported in Fig. 3.

Pearson correlation coefficient ($r$) was also used to assess the correlation between BMD, T-score and Z-score values calculated by the two diagnostic techniques.

3. Results and discussion

For 148 out of the 173 analysed patients, corresponding to 85.55%, US diagnosis (osteoporotic, osteopenic, healthy) coincided with the corresponding DXA one, as visually emphasized by the graphs reported in Figs. 4 and 5.

Fig. 6 shows the corresponding graph obtained for Z-score values. In this case, taking into account the definition of Z-score and the operational definition of osteoporosis, it is not possible the direct identification on the graph of correctly diagnosed patients, false negatives and false positives employing the scheme shown in Fig. 3. However, a statistically significant correlation between US output and corresponding DXA parameter values was found also for Z-score ($r = 0.68, p < 0.001$).

Overall, the diagnostic accuracy of the adopted algorithm, as summarized in Table 1, resulted only slightly inferior to the one recently reported for the same method applied on lumbar spine [37], therefore documenting that the proposed approach can be effectively employed for reliable and non-ionizing osteoporosis diagnoses on central reference sites (i.e., lumbar vertebrae and femoral neck). In fact, the differences in diagnostic accuracy with respect to previously reported results [37] can be attributed to the different size of the enrolled study population (173 patients in the present study, 79 in the previous one) and to the wider considered age range (56–75 y vs 51–60 y).

From data reported in Table 1, it is evident that the maximum diagnostic accuracy (88.00%) was found in correspondence of the youngest investigated patients (56–60 y), while the minimum accuracy (78.57%) was obtained for the oldest recruited women (71–75 y). Therefore, we can say that the adopted algorithm showed a good diagnostic agreement with DXA outcomes for the whole studied age interval, but, on the other hand, a slight effect of patient age on diagnosis accuracy was present and will deserve some further investigations in order to be clarified. However, on the basis of presently available data, we can hypothesize that the observed “trend” is simply due to the fact that the youngest and the oldest age categories were also the less numerous and, consequently, the corresponding results are somehow less reliable. In fact, the other two considered age ranges, which were both much more numerous, showed essentially the same level of diagnostic agreement with DXA evaluations.

For the sake of completeness, the effect of patient age on diagnostic performance was studied also by analysing the correlation between BMD, T-score and Z-score values calculated by the two diagnostic techniques.
coefficient values between single DXA-measured parameters and corresponding US results for each considered 5-year age. The obtained results are reported in Fig. 7.

Actually, the Pearson correlation coefficient between DXA-measured parameters and corresponding US-obtained results showed a somewhat different trend with respect to the discussed behaviour of diagnostic accuracy, and the observed r value trend was roughly reproducible for the three considered diagnostic parameters. In fact, the youngest patients (56–60 y) evidently

**Table 1**

<table>
<thead>
<tr>
<th>Age range (y)</th>
<th>Number of enrolled patients</th>
<th>Number of coincident diagnoses</th>
<th>Diagnostic agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>56–60</td>
<td>25</td>
<td>22</td>
<td>88.00</td>
</tr>
<tr>
<td>61–65</td>
<td>65</td>
<td>56</td>
<td>86.15</td>
</tr>
<tr>
<td>66–70</td>
<td>69</td>
<td>59</td>
<td>85.51</td>
</tr>
<tr>
<td>71–75</td>
<td>14</td>
<td>11</td>
<td>78.57</td>
</tr>
<tr>
<td>Total</td>
<td>173</td>
<td>148</td>
<td>85.55</td>
</tr>
</tbody>
</table>
showed the strongest correlation between DXA diagnostic parameters and corresponding US values (r = 0.85 for BMD, r = 0.84 for T-score, and r = 0.82 for Z-score), while in the other age intervals all the r values were quite close to each other without clearly visible trends (all these r values were in the range 0.68–0.74).

The different trends of diagnostic accuracy and correlation measurements could be probably due to the fact that US measurements are intrinsically affected by bone quality properties, which are an important determinant of actual bone strength [33], while DXA BMD values directly reflect the calcium content measured in the investigated region. This in principle represents an added value of the adopted US approach, since it could integrate bone quantity and bone quality providing a final output that is more closely related to the real bone strength, but further dedicated studies are needed to investigate these aspects through detailed comparisons with gold standard techniques for bone quality assessment (e.g., quantitative computed tomography, micro-indentation, etc.).

Nevertheless, the fact that the highest diagnostic agreement with DXA and the best correlation with single parameters were all found in correspondence of the youngest enrolled patients provides the proposed approach with a specific interesting potential to be employed for mass screenings on young populations.

Referring to papers published by different research groups, the best correlations between US parameters measured at the proximal femur and site-matched BMD measurements were those reported by Barkmann et al. [38]: they found r = 0.85, which is higher than our corresponding result (r = 0.73) but it was also obtained in a significantly smaller patient population (62 vs 173 patients). Furthermore, our backscatter approach was advantageous with respect to the “through transmission” measurements employed in [38] in terms of both bulkiness and complexity of the adopted device.

On the other hand, to the best of our knowledge, the most powerful literature-available results based on a backscatter approach different from the one we proposed in this study were published by Karjalainen et al. [39], who found r = 0.52 between US and DXA evaluations in 26 patients.

Current commercially-available US devices for osteoporosis diagnosis can investigate only peripheral bone districts and present extremely variables degrees of correlation with reference measurements on lumbar vertebrae or femoral neck [40–48]. Therefore, their clinical usefulness is restricted to fragility fracture prediction in patients older than 65 y through calcaneal measurements combined with a detailed assessment of clinical risk factors [49].

As a result, osteoporosis diagnosis is still essentially based on DXA examinations, with impressive evidences of underdiagnosis and undertreatment of this pathology [50,51].

Our proposed approach, which is applicable on the reference axial sites, peculiarly exploits the native integration of the processing of B-mode echographic images and unfiltered RF signals, which is in turn combined with advanced statistical analyses facilitated by the employment of a convex array probe in place of the single-element US sensors typically used in the reported studies. The clinical adoption of our described method for bone densitometry would result in important improvements in osteoporosis management, in particular for what concerns diagnostic test accessibility, thanks to the absence of ionizing radiation use and the possibility of employing the device in primary care settings, like family doctor offices and pharmacies.

4. Conclusion

The clinical feasibility of a novel US-based approach for femoral neck densitometry was demonstrated in a cohort of female patients aged in 56–75 years.

The average diagnostic agreement with the reference gold standard represented by DXA resulted pretty good (85.55%), with a maximum (88.00%) in correspondence of the youngest investigated patients (56–60 y). Overall, diagnostic accuracy showed only minimal variations as a function of patient age, indicating that the proposed approach has the potential to be effectively employable for osteoporosis diagnosis on femoral neck in the whole considered age interval.

Nevertheless, further studies are needed both in order to better quantify the achievable diagnostic performance in the younger and the older populations and also to clarify the nature of single discrepancies with respect to DXA evaluations. Actually, the latter objective will require the employment of additional gold standard references (e.g., quantitative computed tomography), capable of documenting in what measure the outcomes of the proposed US methodology are correlated with bone quality parameters and, therefore, with the actual bone strength even better than DXA.

Acknowledgments

This work was partially funded by FESR P.O. Apulia Region 2007–2013 – Action 1.2.4 (grant n. 3Q5AX31: ECHOLIGHT Project).

References


