Ultrasonics 94 (2019) 109–116

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras

Can low-frequency guided waves at the tibia paired with machine learning differentiate between healthy and osteopenic/osteoporotic subjects? A pilot study

Florian Vogl[a,*], Bernd Friesenbichler[b], Laura Hüsken[a], Inès A. Kramers-de Quervain[c], William R. Taylor[a]

[a] Institute for Biomechanics, ETH Zürich, Zürich, Switzerland
[b] Human Performance Lab, Schulthess Clinic, Zürich, Switzerland
[c] Department for Rheumatology and Rehabilitation, Schulthess Clinic, Zürich, Switzerland

ARTICLE INFO

Keywords:
Osteoporosis
Acoustic waves
Speed of sound
Quantitative acoustics

ABSTRACT

Purpose: Axial transmission quantitative acoustics (ax-QA) has shown to be a promising tool for assessing bone health and properties in a safe, inexpensive, and portable manner. This study investigated the efficacy of low-frequency ax-QA measured at the tibia, paired with a support vector machine (SVM) approach for combining multiple acoustic indicators, to diagnose osteoporosis as defined by bone mineral density.
Methods: This pilot study measured 41 female subjects using ax-QA (flexural mode, 3 kHz) at the tibia and using dual X-ray absorptiometry (DXA) at the lumbar spine, femoral neck, and distal radius. For each location, a threshold classifier and SVM were trained to differentiate between healthy and non-healthy subjects based on the phase velocity at different frequencies. Receiver Operating Characteristics and area under curve values (AUC) were used to assess the classifiers' performances for various thresholds and class-weights.
Results: The SVM outperformed the threshold classifier for all three bone locations at low false positive rates. While differentiation between healthy and non-healthy bone states was poor for the spine (AUC: 0.56 ± 0.04), good to moderate performances were observed for the radius (AUC: 0.83 ± 0.03) and hip (AUC: 0.71 ± 0.04).
Conclusions: Low-frequency ax-QA has demonstrated potential for complementing DXA in screening for osteoporosis at the radius and hip. Through further addition of acoustic indicators ax-QA could provide a diagnostic alternative in third-world countries, and bring bone health screening and monitoring into the hands of clinicians and general health practitioners everywhere.

1. Introduction

Osteoporosis is the world's most prevalent skeletal disorder, affecting some 50% of women and 20% of men over the age of 50 in the western world, and more than 200 million individuals worldwide [1,2]. Osteoporosis is characterised by a deterioration of bone properties at the microscopic and macroscopic levels [1,3] and is clinically diagnosed when the T-score of the bone mass density (BMD) at the lumbar spine, the distal radius, or the femoral neck falls below −2.5 or when a fracture occurs with inadequate trauma. The associated deterioration of skeletal mechanical and structural properties leads to an increased osteoporotic fracture risk, as well as increased accompanying rates of mortality [4] and morbidity [5], often necessitating long-term care or assistance with fundamental daily tasks [6]. It is estimated that one in two women and one in five men over the age of 50 will suffer from an osteoporotic fracture [7] – numbers that will likely be exacerbated by increasing life expectancies. Osteoporosis thereby poses a serious economic challenge, with the total cost of osteoporotic fractures having been estimated to be about $8 billion per year in Australia alone when considering all direct and indirect costs [8].

Despite the apparent importance of addressing osteoporosis in a comprehensive and efficient manner, the clinical diagnosis and treatment of osteoporosis remains inadequate, with between 75% and 90% of high-risk individuals estimated to remain unevaluated [8,9]. One key contributing factor to this chronic under-assessment could be due to the current gold standard for BMD assessment: Dual X-ray Absorptiometry (DXA). While this technique has shown efficacy in clinical practice, DXA suffers from multiple drawbacks, including (a) exposure to

* Corresponding author at: ETH Zürich, Institut für Biomechanik, Leopold-Ruzicka-Weg 4, 8093 Zürich, Switzerland.
E-mail address: florian.vogl@hest.ethz.ch (F. Vogl).

https://doi.org/10.1016/j.ultras.2018.11.012
Received 20 December 2017; Received in revised form 4 July 2018; Accepted 29 November 2018
Available online 30 November 2018
0041-624X/

ionizing-radiation and its related risks, (b) high equipment costs, (c) limited portability, and (d) the requirement for trained operators. The measurement of the BMD alone lacks information on structural- (e.g. cortical thickness) [10] and material properties (e.g. anisotropy) of the bone and does not fully explain the fracture risk of a subject. In fact, a majority of low-trauma fractures occur in non-osteoporotic subjects with only moderately decreased BMD [11]. Moreover, widespread screening for osteoporosis is generally not undertaken without specific clinical indication, resulting in osteopenia and osteoporosis often being diagnosed only when a fracture occurs.

One promising approach for widespread screening is quantitative acoustics (QA), also commonly referred to as quantitative ultrasound when frequencies above 20 kHz are employed [12]. QA is a method for bone health assessment based on measuring the propagation of acoustic waves within the bone. In addition to being radiation-free, the technique is inexpensive, portable, and easy-to-use, thereby making QA a prime candidate for screening and widespread monitoring applications. Most work in the QA field has focused on acquiring structural and material information that is unavailable using radiative methods, with the goal of complementing or potentially even outperforming the current radiative gold standards [13–19]. Other researchers have investigated possibilities of using QA to obtain equivalent information to current radiation-based methods [20–25], which, due to the unique advantages of QA, could bring bone health monitoring into the hands of clinicians and general health practitioners as a component of routine screening. As a consequence, the development of QA techniques to diagnose and screen for osteoporosis could help address the aforementioned problem of under-diagnosis within current clinical practice.

Towards this end, several variants of QA for bone assessment are currently being investigated, including scanning acoustic microscopy [26], through-transmission [27–29], backscatter [30,31], axial- or circumferential transmission techniques [16,18,32–34]. While these techniques are all based on the measurement of acoustic waves, the methods vary considerably in underlying measurement principle and specific advantages. Axial transmission quantitative acoustics (ax-QA) is a QA technique for the assessment of cortical bone, in which transducers are placed along the bone and transmit acoustic waves into the cortex through the overlying soft-tissues. Separate surface sensors then measure the propagation of these waves, from which information on bone structural and material properties can be derived. Advantages of this technique include the requirement for only unilateral access to the bone and its ability to characterise a multitude of properties over a large region and cross-section of the bone, even within a single measurement. It is important to note that due to the dispersive nature of human long bones, the choice of excitation frequency and assessed wave mode(s) governs which mechanical bone properties can be measured [14,32,35–37].

Ax-QA approaches are often based on analyzing the fundamental symmetrical S0 wave mode, which can often be identified with the first-arriving signal (FAS), or the assessment of the fundamental antisymmetric A0 wave mode, often identified as the energetic late arrival (ELA). FAS-based approaches have exhibited only limited sensitivity to DXA-defined osteoporosis [38–40], even though they have demonstrated efficacy in fracture discrimination [41–43]. Compared to measurements of the S0-mode, measurements of the A0-mode are known to be sensitive to changes in bone properties near the endosteum, including porosity [37,44], BMD [45], bending stiffness [10], and cortical thickness, especially at low frequencies below about 300 kHz [17,32,37]. The assessment of these properties is critical for the identification of osteoporosis [46–49], thus making low-frequency assessment of the A0-mode a promising approach for osteoporosis diagnosis and monitoring.

Multiple studies have shown that QA approaches that differ in fundamental principle, excitation frequency, or measurement location, are able to yield complementary information on the underlying bone properties [32,42,50,51]. It is therefore desirable to develop scalable analysis approaches that are able to efficiently incorporate various sources of acoustic and clinical information for each specific assessment. Previous approaches to combine such multi-faceted information have been based on statistical models, such as logistic regression [42] or linear combinations of multiple features [52], or based on physical models, such as multi-mode curve fitting of a wave guide model to the experimental dispersion relation [14]. While these approaches are powerful, especially for the determination of bone properties, they are all restricted by the need to explicitly specify the underlying model, as well as the limitations imposed by the model itself.

For osteoporosis detection, we propose an approach based on a support vector machine (SVM) [53,54]. A SVM is a discriminative supervised learning algorithm for regression or classification that is itself able to determine the underlying model directly from the data, and has successfully been applied to a variety of technological and scientific applications, including handwriting recognition [55], ultrasonic flaw detection in non-destructive testing [56], and ultrasonic determination of bone demineralization [57] The SVM's sparse kernel-based formulation allows for a rich class of models and permits efficient application to high-dimensional feature spaces, making SVMs potentially suitable as a scalable, flexible, and extendable technique for osteoporosis detection in clinical settings.

The aim of the presented work was therefore to investigate the potential of low frequency A0-mode ax-QA, combined with a SVM-based analysis approach, to diagnose DXA-defined osteoporosis.

2. Methods

In this study 41 female subjects were measured using ax-QA at the tibia and using dual X-ray absorptiometry (DXA) at the lumbar spine, femoral neck, and distal radius. Based on the acoustic data a threshold classifier and SVM were trained for each location to differentiate between healthy and non-healthy (osteoporotic/osteopenic) subjects, with DXA results being considered the ground truth. Receiver Operating Characteristics and area under curve values (AUC) were used to assess the classifiers' performances for various thresholds and class-weights.

2.1. Cohort

Eligible subjects had to be female, as women are more than twice as likely to suffer from osteoporosis as men [2]. Furthermore, eligible subjects had to be between 20 and 80 years old, and had to have DXA-information available from a DXA-examination at the investigation site within the previous year, acquired as part of clinical routine. Mean age of the participants and standard deviation were 61.1 ± 9.6 years. Exclusion criteria were: fracture of the tibia, knee, or foot within the previous year; a prosthesis or implant at the tibia, knee, or foot; a rash, wound, or infection at the tibia. All measurements were conducted at the Schulthess Clinic, Zurich, Switzerland. Ethical approval for this study was granted by the ethics committee of ETH Zurich under the reference: EK 2014-N-57 and informed consent was obtained from all individual participants included in the study.

2.2. Dual X-ray absorptiometry (DXA)

DXA measurements were taken at the clinical standard locations of femoral neck, distal radius, and lumbar spine (L1-L4) using a "Horizon A" DXA device (Hologic, Massachusetts, USA), which was periodically calibrated as part of clinical routine. For two subjects, DXA values for the distal radius were not available, because of one recent fracture and one implant in this region of interest. Data from these two subjects were excluded from any analysis involving the radius region, but were included in all other analyses.

From the BMD-value $BMD_{s,l}$ of each subject s and each location l the T-score was calculated as:

Table 1
Constants used in the determination of T-scores at the clinical standard locations.

Region of interest	PPMBMD	PPMSTD	Source
Lumbar spine	1.018	0.110	Hologic Database
Femoral neck	0.942	0.122	Hologic Database (NHANES)
Distal radius	0.564	0.051	Hologic Database

Table 2
Number of subjects for each of the three bone locations, categorised into the three bone health states based on BMD T-score.

	Distal radius	Femoral neck	Lumbar spine
Healthy	14	18	15
Osteopenic	15	21	20
Osteoporotic	9	1	5

$$T_{s,l} = BMD_{s,l} - (PPMBMD_l/\mathrm{PPMSTD_l})$$

where *PPMBMD* is the peak population mean BMD and *PPMSTD* is its standard deviation (Table 1).

The T-scores were then categorised into bone health states according to the World Health Organisation's definitions: healthy for $T > 0$, osteopenic for $0 > T > -2.5$, and osteoporotic for $T < -2.5$. Table 2 lists the number of subjects for each health status for each of the bone locations.

2.3. Acoustic measurements

Forty subjects were measured successfully using ax-QA at the tibia. Acquisition of acoustic signals failed for one subject because the linear sensor array could not be lined up with the subject's unusually curved tibiae. Consequently, data from this subject were excluded from all analyses. While details on the measurement protocol, measurement device, analysis algorithm and their reliabilities are all described in a previous work [58], the following sections present a brief overview. Please note that the analysis algorithm did not contain any free parameters and was not adjusted in any way to the current study.

2.3.1. Device

All acoustic measurements in the study were performed using the Bone Stiffness Measurement Device (BSMD), which consists of three main components: a piezo-electric transducer to excite an acoustic wave, four acceleration sensors to measure the wave propagation, and a data acquisition system to control the device and record the measurement data. The excitation, sent to the transducer, was a sine waveform of 3000 Hz with a Gaussian envelope of Full-Width-at-Half-Maximum of 1000 Hz. The four acceleration sensors were mounted along the tibial axis in a linear array with 2 cm inter-sensor distance. The sensors had a sensitivity of 10 mV/G ± 0.1 mV/G, where G is the gravitation constant. 200 data points per sensor were collected per acoustic measurement at a sampling rate of 96 kHz.

2.3.2. Measurement protocol

At the beginning of each measurement, the operator palpated the malleolus medialis of the left leg as well as the joint gap above the left medial tibial condyle and marked the mid-point between these two landmarks along the tibial facies medialis. The sensor array was centred on the marked position and fixed using two straps. The transducer was positioned on the left tibial head and the excitation pulse was applied, creating a flexural wave (A0-like mode from Lamb plate theory) inside the bone. The propagation of this wave was measured using the acceleration sensors, digitized by the data acquisition system, and transferred to the laptop for off-line analysis (Fig. 1). Fifty such wave measurements were taken in rapid succession before transferring the transducer and sensor array to the right leg. Following this procedure, both legs were measured alternately three times each, yielding a total of 150 wave measurements per tibia for each subject.

2.3.3. Phase velocity analysis

Each acoustic measurement was composed of four channels, corresponding to the four acceleration sensors (Fig. 1). For each channel, the signal was low-pass filtered using a tenth-order Butterworth filter in a zero-phase configuration with a cut-off frequency of 8 kHz. The signals were then windowed using Gaussian windows with a standard deviation of 14 samples, while the mode of each window was set to the midpoint between the first valley and the following peak, where a valley and peak were defined as exceeding 1/3 of the signal minimum or maximum respectively. Measurements with a midpoint prior to sample 40 were considered invalid due to premature triggering of the data acquisition and were excluded from further analysis.

After windowing, each channel was transformed from the time into the frequency domain using a Fast-Fourier-Transform of length 2048 points and split into magnitude and phase. The phases were then unwrapped around the centre frequency of 3000 Hz and phase differences $\Delta\phi$ between the channels $i, j \in \{0, 1, 2, 3\}$ were calculated as:

$$\Delta\phi_{i,j} = \phi_i - \phi_j.$$

Since the phase was expected to change strongly between sensors but only weakly as a function of frequency, measurements with phase crossings were considered invalid and discarded from further analysis. In total, around 15% of all measurements were discarded due to phase crossings or premature triggering (Table 2).

From the phase differences, the phase velocity v between the channels $i, j \in \{0, 1, 2, 3\}$ was determined from

$$v_{i,j} = v_{i,j}(f) = \frac{\Delta s_{i,j}\, 2\pi \mathrm{f}}{\Delta\phi_{i,j}},$$

where f is the frequency and $\Delta s_{i,j}$ is the distance between the sensors corresponding to channel i and channel j. The arithmetic mean over these pairwise phase velocities resulted in the overall phase velocity value for the acoustic measurement.

2.4. Statistics

The presented work quantified the efficacy of two classifiers, one simple threshold classifier and one support vector machine, to correctly diagnose the BMD-based health status for the three bone locations from the acoustic measurements. Due to the low number of patients classified as osteoporotic in the clinical testing (Table 2), the categories "osteopenic" and "osteoporotic" were combined into a "non-healthy" class, thus creating three binary classification tasks with imbalance ratios, calculated as #measurements (non-healthy)/#measurements (healthy), of 1.8 for the radius, 1.2 for the femoral neck, and 1.8 for the lumbar spine.

(a) Threshold classifier

The threshold classifier used the phase velocity at the central frequency of 3000 Hz as the only input feature, and classified each measurement according to whether this phase velocity was above or below a certain threshold.

(b) Support vector machine

This SVM classifier used the phase velocity values at seven different frequencies as input features: 2719, 2812, 2906, 3000, 3094, 3188, and 3281 Hz. These frequencies were chosen centered on the experimental excitation central frequency and as a consequence of the chosen sampling frequency and Fast-Fourier-Transform length (Fig. 1). For these features, a soft-margin SVM using a radial Gaussian kernel was trained, the implementation of which was based on the "scikit-learn" library [59]. A 10×6 nested cross-validation design was employed to perform hyperparameter optimisation during training, and to estimate the

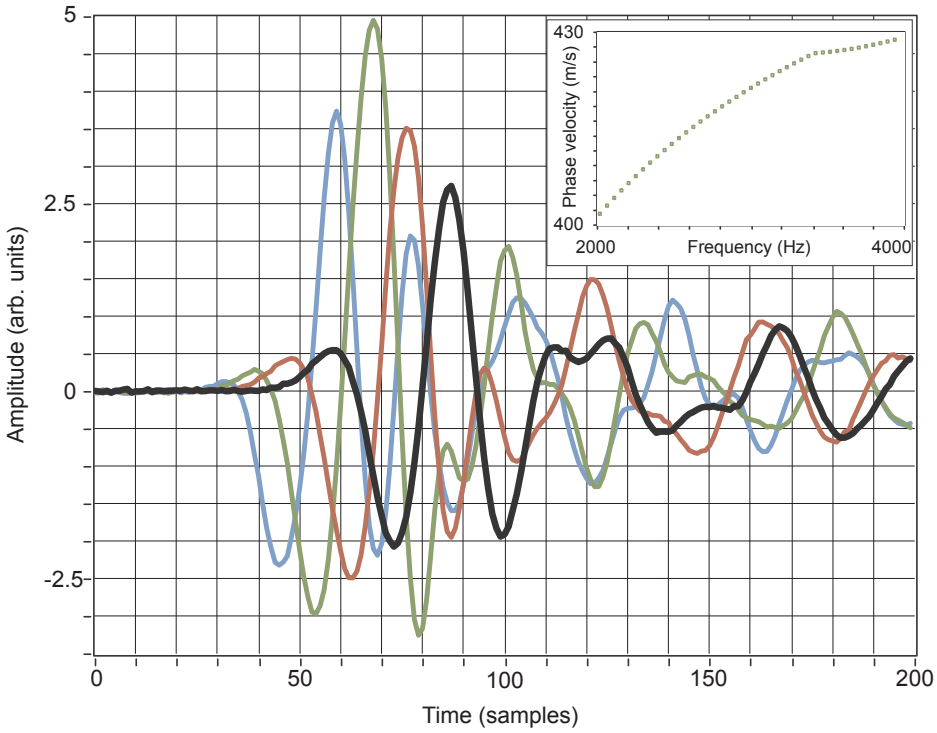

Fig. 1. Example of an acoustic measurement: the main figure shows the four accelerations measured by the sensors as a function of time, and the inset depicts the derived phase velocity as a function of frequency.

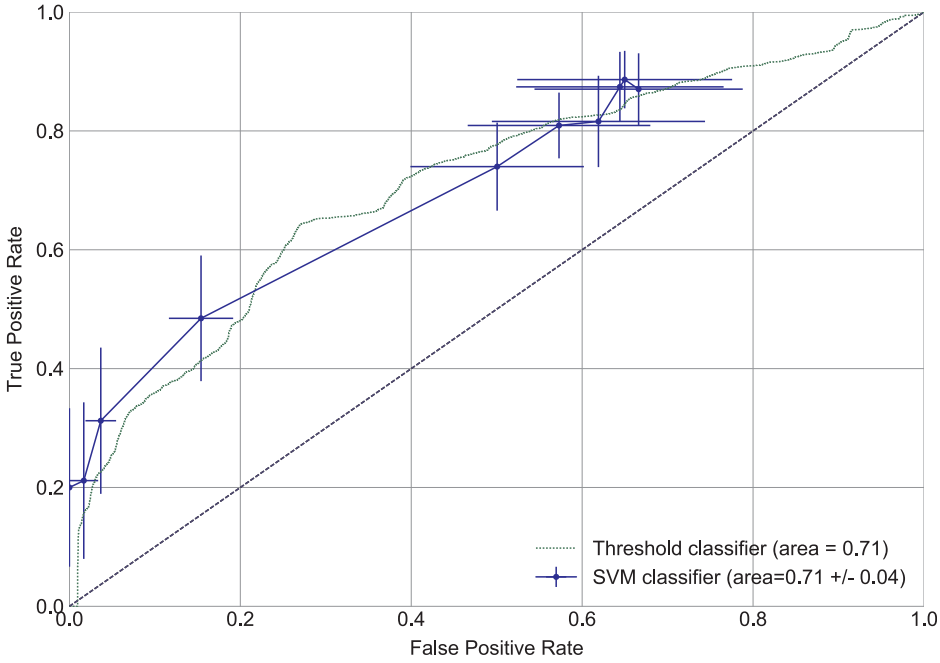

Fig. 2. Receiver-operating characteristic (mean ± standard error) for classification of healthy versus non-healthy bone state at the femoral neck. For the SVM the points correspond to different values for the classweight of the healthy class: 0.0001, 0.01, 0.1, 1.0, 2.0, 3.0, 5.0, 7.0, 50.0, and 1000.0 (from left to right).

classifier's final generalised performance and its uncertainty – i.e. how well the SVM approach would perform if applied to the underlying population [20]. The outer cross-validation used subject grouping, avoid overfitting to subject-specific features by ensuring that each subject's data is either part of the training or test-set for each fold, but never both. The inner cross-validation used randomized parameter optimization with exponential distributions to optimize the Gaussian kernel variance and the regularisation parameter for each of the outer folds. From the confusion matrix of the test set for each fold, the true positive rate (TPR), false positive rate (FPR), true negative rate (TNR) and false negative rate (FNR) were calculated. The arithmetic mean and its standard error over the individual folds was used to estimate the classifier's generalised performance and its uncertainty. This procedure was executed for ten different class weights of the healthy bone class (0.0001, 0.01, 0.1, 1.0, 2.0, 3.0, 5.0, 7.0, 50.0, and 1000.0) in order to investigate the performance of the classifier in situations where it is more important to correctly classify one class than misclassify the other.

Standard receiver operating characteristics (ROC) and the area

under curve (AUC) measure were used to visualise and quantify the performance of the classifiers. The AUC for both classification methods was estimated using a trapezoidal rule summation:

$$\mathrm{AUC} = \sum_i \frac{y_i + y_{i+1}}{2}(x_{i+1} - x_i),$$

which is a conservative estimate for concave underlying functions. Here, x_i, y_i denote the FPR and TPR of classifier instance i with corresponding parameters $\{p_i\}$, for example the FPR and TPR of the threshold classifier with a threshold value of 400 m/s. The points (0,0) and (1,1) in ROC-space were included for the purpose of calculating the AUC. For the SVM classifier, the uncertainty of the AUC was estimated by Gaussian error propagation as

$$\Delta AUC = \frac{1}{2}\sqrt{\sum (y_{i-1} - y_{i+1})^2 \Delta x_i^2 + (x_{i-1} - x_{i+1})^2 \Delta y_i^2},$$

where Δx_i and Δy_i are the aforementioned standard errors of the TPR and FPR for classifier instance i.

3. Results

The AUC values of the threshold classifier were 0.71 for the hip (Fig. 2), 0.81 for the radius (Fig. 3), and 0.58 for the spine (Fig. 4), while the corresponding SVM classifier values were 0.71 ± 0.04, 0.83 ± 0.03, and 0.56 ± 0.04. Although the measurement of low-frequency guided waves at the tibia was only able to slightly outperform random chance for the classification of bone health at the spine, this approach exhibited good to moderate performance for differentiating between healthy and non-healthy (osteopenic/osteoporotic) bone states at the distal radius and the hip.

4. Discussion

The presented work investigated the suitability of ax-QA measured at the tibia for diagnosing osteoporosis, with the aim to bring bone health monitoring into the hands of clinicians and general health practitioners to facilitate widespread routine screening, and help to address the problem of under-diagnosis. Moreover, this low-cost and portable approach could provide an important diagnostic alternative in e.g. under-developed countries, where larger populations have limited access to expensive clinical equipment.

While not directly comparable, the observed differences in classification performance between target locations are consistent with previous studies: one study found Pearson correlation coefficients of 0.31–0.47 between tibial speed of sound (SoS) and femoral neck BMD, and 0.63 between tibial SoS and radial BMD [60]; Cohen's Kappa values of 0.29, 0.27 and 0.46 to detect BMD T-values below $T < -2.5$ at the lumbar spine, femoral neck and radius [39]; sensitivity/specificity of 0.5/0.8 and 0.6/0.65 to detect BMD T-values below $T < -1$ at the lumbar spine and at the femoral neck [38], and; correlation coefficients of 0.29 between tibial SoS and lumbar spine BMD, and 0.36 between tibial SoS and femoral neck BMD [38]. Here, the consistently higher prediction performance of radius values over hip values from measurements at the tibia is somewhat surprising, considering that skeletal adaptation is thought to be governed by the internal mechanical loading conditions, which are plausibly most similar throughout the lower limb [61]. The facts, that ax-QA is known to be sensitive to BMD [62] and that the relationship between BMD values of the tibia and the lumbar spine is known to be low [63], could explain the low performance for predicting spinal bone state. However, there seems to be no clear consensus on the relative strength of the relationships between tibial BMD and radial BMD versus tibial BMD and hip BMD [63,64]. Even though it seems clear that the differences between target locations observed in all these studies are partially due to the sensitivity of ax-QA to BMD paired with the known relationships between BMD locations, further effects that are not captured by BMD must be present. It is likely, that these effects relate to differences in the biomechanical structure, where the geometry and bone composition of the tibia is more similar to the radius (e.g. primarily cortical bone) than to the femoral neck and lumbar spine (e.g. primarily trabecular bone). Additionally, spinal BMD values are known to be frequently distorted due to degenerative deformities common in the study population, potentially rendering the correct classification of bone health status at the spine more challenging than at the other two bone sites.

The AUC value for the hip of 0.71 ± 0.04 is similar to the findings of another study measuring ax-QA at the tibia in 93 subjects [22]. Their study combined ax-QA measurements at frequencies of 0.1 MHz and

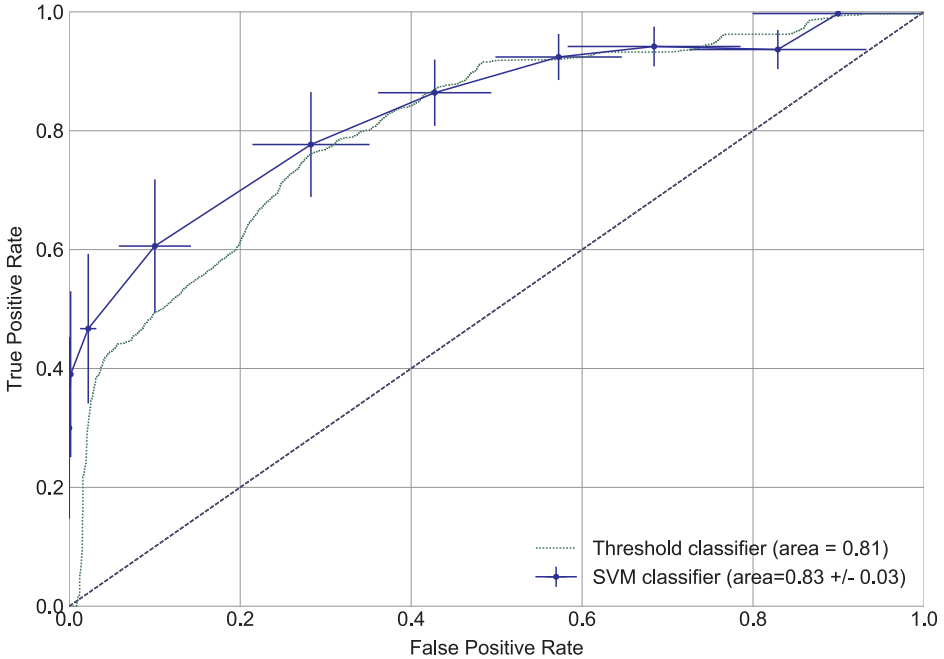

Fig. 3. Receiver-operating characteristic (mean ± standard error) for classification of healthy versus non-healthy bone state at the distal radius. For the SVM the points correspond to different values for the classweight of the healthy class: 0.0001, 0.01, 0.1, 1.0, 2.0, 3.0, 5.0, 7.0, 50.0, and 1000.0 (from left to right).

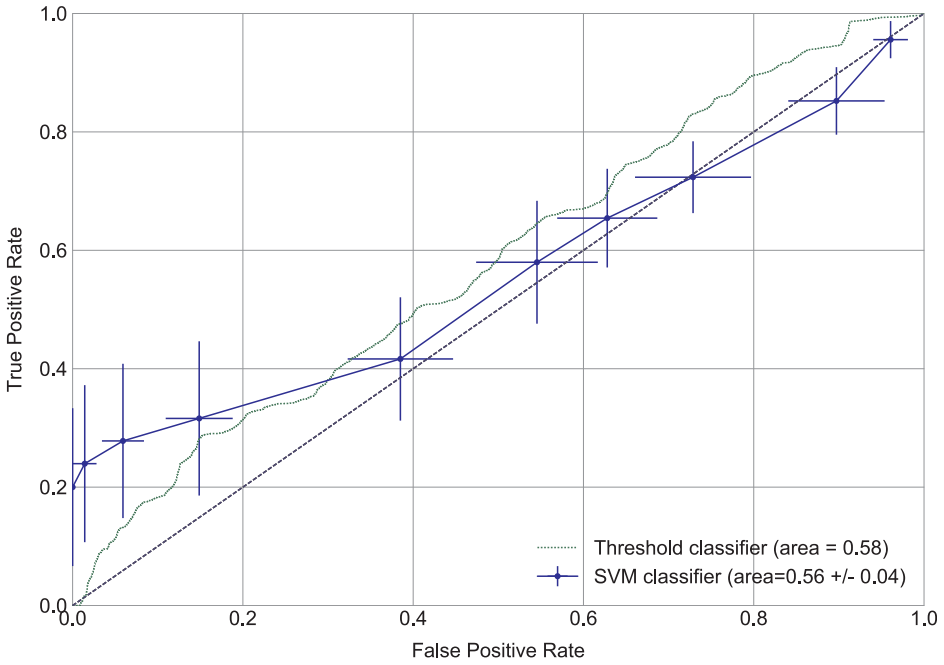

Fig 4. Receiver-operating characteristic (mean ± standard error) for classification of healthy versus non-healthy bone state at the lumbar spine. For the SVM the points correspond to different values for the classweight of the healthy class: 0.0001, 0.01, 0.1, 1.0, 2.0, 3.0, 5.0, 7.0, 50.0, and 1000.0 (from left to right).

1 MHz to discriminate between healthy and osteopenic/osteoporotic BMD T-scores at the hip and reported an AUC value of 0.79, as well as a sensitivity and specificity of 76% and 70%, corresponding to a TPR and FPR of 0.76 and 0.3. Potential reasons for the slightly higher performance compared to the results of the current study could include (a) employment of different acoustic features; (b) the choice of examination frequency, specifically 100 kHz compared to the 3 kHz used in our study; (c) the inclusion of additional high-frequency information at 1 MHz; and d) overestimation of the performance due to overfitting of the multi-parametric model [20].

Another study on 331 subjects measured ax-QA at the tibia using multiple frequencies between 60 kHz and 1200 kHz [21]. From the ax-QA measurements, ten acoustic parameters were extracted and evaluated for their ability to discriminate whether the subjects' lumbar BMD T-scores were above or below −2.5. The reported sensitivity and specificity for their approach were 87% and 63%, corresponding to a TPR of 0.87 and FPR of 0.37. Similar values of AUC = 0.6–0.7 and TPR/FPR of 0.79/0.42 have been reported for the detection of combined spine and hip BMD T-values below $T < -2.5$ from multi-site FAS measurements [65] or calcaneal transmission QA [66]. It is important to note, however, that these studies all discriminated between the osteoporotic class ($T < -2.5$) and a merged osteopenic and healthy class ($T > -2.5$), while our study discriminated between the healthy class ($T > -1$) and a non-healthy class with $T < -1$. Importantly, however, the difference in binary classification outcome made direct comparison of the relative performances difficult.

While cross-validation allowed the estimation of generalised accuracy and its uncertainty from our sample, our pilot study was constrained by the unexpectedly low number of subjects with osteoporotic BMD values, especially at the spine and hip, necessitating two of the bone states to be merged in our study. We attribute the low number osteoporotic participants to a mixture of sampling variability and possible selection-bias: as subjects were not blinded to their medical condition, osteoporotic patients might have been less inclined to participate in the study. After considering which of the classes to merge, we chose to differentiate between healthy and osteopenic/osteoporotic bone states, due to the fact that a high number of fractures occur in osteopenic subjects [11,67]. Consequently, such a non-invasive approach could allow an inexpensive clinical screening tool for identifying high-risk subjects and thus improve medical resource allocation.

Overall, although the differences in measurement location, excitation frequency, analysed wave mode, analysis method, and class definition make direct comparisons challenging, our ax-QA approach performed on a level similar to other results presented in the literature for the identification of osteoporotic/osteopenic subjects. However, due to the low number of osteoporotic subjects in our study, the presented results for the binary differentiation are likely to underestimate the true performance of our approach for the hip and spine. A follow-up study on a larger cohort with an even distribution of bone health states will be required to confirm our results and investigate whether our approach can also differentiate between all three bone states.

It is also noteworthy that the acoustic examination location in our study (tibia) did not correspond to the diagnostic regions of interest for the DXA examinations (distal radius, femoral neck, lumbar spine). For FAS-based ax-QA approaches, at least, correlations between the SoS and BMD have shown to be site-dependent, with highest values typically found for site-matched acoustic measurements [38,60,68]. This site-dependency suggests that ax-QA measurements at different locations could yield complementary information, much like the synergistic information offered by using different wave modes and frequencies [14,32,35–37]. Even though further work is needed to investigate the advantages of each QA-variant, the proposed SVM approach clearly shows potential for directly combining information from measurements at different locations, using multiple frequencies and various modes to enhance classification performance. As such, we personally hope that the question of which frequency range or QA-variant is "better" should be supplanted by: how can we successfully combine acoustic approaches and characteristics towards specific applications? Here, we believe that the proposed SVM approach might be one way of achieving this goal in the future.

Although adaption of ax-QA to the radius is straight-forward and is currently one of the standard measurement locations under investigation, in-vivo application of ax-QA to the hip is still complicated by the overlying soft-tissues. In addition, application of ax-QA at both the spine and hip are complicated by the geometric deviation from tubular geometry, which, while not a fundamental limitation, is an assumption

made by most ax-QA approaches. The extension of current ax-QA techniques to include multiple frequencies and modes is currently under active investigation and might offer novel approaches for addressing such issues [14–16,21].

The combination of multiple information sources towards a specific application could be performed using an SVM approach similar to the technique presented in our study, which was able to successfully combine the information from multiple frequencies. Importantly, the SVM outperformed the simple threshold classifier for all three bone locations in the regions of low FPR (Figs. 2–4), which is the region most relevant for the majority of clinical applications. Of course, the additional information available to the SVM over the threshold classifier in our study was limited to a narrow frequency range around the excitation frequency, leading to only a modest improvement in classification performance. Whether this SVM approach is also capable of combining more dissimilar or disparate measures remains to be confirmed. One limitation of the SVM is that it produces outputs as decisions rather than as posterior probabilities. Such posterior probabilities could be useful in clinical situations in which the results of multiple tests should be combined, prior information needs to be considered, or a flexible trade-off between sensitivity and specificity is needed. Here, an alternative might be a relevance vector machine (RVM) [69], which is a Bayesian kernel technique sharing many commonalities with the SVM approach. The drawback for the probabilistic output is mainly the RVM's longer training times compared to the SVM, which is partly offset because no cross-validation is needed to find the optimal model parameters.

5. Conclusion

Low-frequency axial transmission quantitative acoustics has shown to be a potential method capable of complementing DXA in the screening of osteoporosis. A clear dependency of detection performance on target location was shown, with good to moderate performances observed for the radius and the hip and poor performance for the spine. The SVM approach was shown to be successful in combining phase velocities at different frequencies, thus offering perspectives for future improvements by combining information from multiple measurement locations, frequencies, and wave modes. Ultimately, such a device could provide a diagnostic alternative in areas with limited access to DXA equipment, such as third-world countries, and bring bone health screening and monitoring into the hands of clinicians and general health practitioners everywhere.

Acknowledgments

Funding: This study was partially funded by the RMS Foundation, Bettlach, Switzerland.

Conflicts of interest

Florian Vogl, Bernd Friesenbichler, Laura Hüsken, Inès A. Kramers-de Quervain, and William R. Taylor declare that they have no conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ultras.2018.11.012.

References

[1] J. Lin, J. Lane, Osteoporosis: a review, Clin Orthop Relat Res 8 (2004) 126–134.
[2] T.P. Van Staa, E.M. Dennison, H.G.M. Leufkens, C. Cooper, Epidemiology of fractures in England and Wales, Bone 29 (2001) 517–522, https://doi.org/10.1016/S8756-3282(01)00614-7.
[3] H.R. Fish, R.F. Dons, Primary osteoporosis, Am. Fam. Physician 31 (1985) 216–223, https://doi.org/10.1016/S0140-6736(80)91273-8.
[4] R. Gandhi, D. Tsvetkov, J.R. Davey, N.N. Mahomed, Survival and clinical function of cemented and uncemented prostheses in total knee replacement: a meta-analysis, J. Bone Joint Surg. Br. 91 (2009) 889–895, https://doi.org/10.1302/0301-620X.91B7.21702.
[5] A.G. Randell, T.V. Nguyen, N. Bhalerao, et al., Deterioration in quality of life following hip fracture: a prospective study, Osteoporos. Int. 11 (2000) 460–466.
[6] W. Scaf-Klomp, E. Van Sonderen, R. Sanderman, et al., Recovery of physical function after limb injuries in independent older people living at home, Age Age. 30 (2001) 213–219, https://doi.org/10.1093/ageing/30.3.213.
[7] NIH Consensus Development Panel on Osteoporosis Prevention Diagnosis and Therapy, Osteoporosis prevention, diagnosis, and therapy. JAMA 285 (2001) 785–795.
[8] T.V. Nguyen, J.A. Eisman, Osteoporosis: underrated, underdiagnosed and undertreated, Med. J. Aust. 180 (2004) S18–S22.
[9] J.T. Harrington, S.B. Broy, A.M. Derosa, et al., Hip fracture patients are not treated for osteoporosis: a call to action, Arthritis Rheum. 47 (2002) 651–654, https://doi.org/10.1002/art.10787.
[10] E. Stüssi, H. Bischof, E. Lucchinetti, et al., Development and adaptation of bending stiffness of the skeleton of the extremities as exemplified by the human tibia through exercise, Sport Sport 8 (1994) 103–110.
[11] A. Cranney, S.A. Jamal, J.F. Tsang, et al., Low bone mineral density and fracture burden in postmenopausal women, Can. Med. Assoc. J. 177 (2007) 575–580.
[12] P. Laugier, G. Haïat, Bone quantitative ultrasound, Springer, Dordrecht. doi: 10.1007/978-94-007-0017-8.
[13] P.H.F. Nicholson, P. Moilanen, T. Kärkkäinen, et al., Guided ultrasonic waves in long bones: modelling, experiment and in vivo application, Physiol. Meas. 23 (2002) 755–768.
[14] N. Bochud, Q. Vallet, Y. Bala, et al., Genetic algorithms-based inversion of multimode guided waves for cortical bone characterization. Phys. Med. Biol. 6953:6953. doi: 10.1088/0031-9155/61/19/6953.
[15] N. Bochud, Q. Vallet, J.-G. Minonzio, P. Laugier, Predicting bone strength with ultrasonic guided waves, Sci. Rep. 7 (2017) 43628, https://doi.org/10.1038/srep43628.
[16] Q. Vallet, N. Bochud, C. Chappard, et al., In vivo characterization of cortical bone using guided waves measured by axial transmission, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 63 (2016) 1361–1371, https://doi.org/10.1109/TUFFC.2016.2587079.
[17] P. Moilanen, M. Talmant, P.H.F. Nicholson, et al., Ultrasonically determined thickness of long cortical bones: three-dimensional simulations of in vitro experiments, J. Acoust. Soc. Am. 122 (2007) 2439–2445, https://doi.org/10.1121/1.2769619.
[18] P. Moilanen, M. Määttä, V. Kilappa, et al., Discrimination of fractures by low-frequency axial transmission ultrasound in postmenopausal females, Osteoporos. Int. 24 (2013) 723–730, https://doi.org/10.1007/s00198-012-2022-x.
[19] M.K.H. Malo, D. Rohrbach, H. Isaksson, et al., Longitudinal elastic properties and porosity of cortical bone tissue vary with age in human proximal femur, Bone 53 (2013) 451–458, https://doi.org/10.1016/j.bone.2013.01.015.
[20] R. Kohavi, A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, in: IJCAI'95 Proc. 14th Int. Jt. Conf. Artif. Intell. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995, pp 1137–1143.
[21] V. Egorov, A. Tatarinov, N. Sarvazyan, et al., Osteoporosis detection in postmenopausal women using axial transmission multi-frequency bone ultrasonometer: clinical findings, Ultrasonics 54 (2014) 1170–1177, https://doi.org/10.1016/j.ultras.2013.08.017.
[22] A. Tatarinov, V. Egorov, A. Sarvazyan, The dual-frequency method for ultrasonic assessment of skeletal system, Acoust. Phys. 55 (2009) 665–673, https://doi.org/10.1134/S106377100904023X.
[23] Francesco Conversano, Roberto Franchini, Antonio Greco, F.C. Giulia Soloperto, E. Casciaro, M. Aventaggiato, et al., A novel ultrasound methodology for estimating spine mineral density, Ultrasound Med. Biol. 41 (2015) 281–300, https://doi.org/10.1016/j.ultrasmedbio.2014.08.017.
[24] J.P. Karjalainen, O. Riekkinen, J. Töyräs, et al., New method for point-of-care osteoporosis screening and diagnostics, 2016, 971–977. doi: 10.1007/s00198-015-3387-4.
[25] E.M. Stein, F. Rosete, P. Young, et al., Clinical assessment of the 1/3 radius using a new desktop ultrasonic bone densitometer, Ultrasound Med. Biol. 39 (2013) 388–395, https://doi.org/10.1016/j.ultrasmedbio.2012.09.024.
[26] K. Raum, J. Reisshauer, J. Brandt, Frequency dependence of the anisotropic impedance estimation in cortical bone using time-resolved scanning acoustic microscopy, vol.2, in: 2002 IEEE Ultrason. Symp. 2002. Proceedings, 2002, pp. 1301–1304.
[27] R.B. Ashman, S.C. Cowin, W.C. Van Buskirk, J.C. Rice, A continuous wave technique for the measurement of the elastic properties of cortical bone, J. Biomech. 1 (1984) 349–361.
[28] P.H.F. Nicholson, G. Lowet, C.M. Langton, et al., A comparison of time-domain and frequency-domain approaches to ultrasonic velocity measurement in trabecular bone, Phys. Med. Biol. 41 (1996) 2421–2435.
[29] K. Mizuno, Y. Nagatani, K. Yamashita, M. Matsukawa, Propagation of two longitudinal waves in a cancellous bone with the closed pore boundary, J. Acoust. Soc. Am. 130 (2011), https://doi.org/10.1121/1.3607196 EL122-EL127.
[30] O. Yousefian, Y. Karbalaeisadegh, M.M. Muller, et al., Diffusion constant and frequency dependent attenuation for the assessment of cortical bone porosity, J. Acoust. Soc. Am. 142 (2017) 2565, https://doi.org/10.1121/1.5014380.
[31] B.K. Hoffmeister, D.P. Johnson, J.A. Janeski, et al., Ultrasonic characterization of

human cancellous bone in vitro using three different apparent backscatter parameters in the frequency range 0.6-15.0 mhz, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 55 (2008) 1442–1452, https://doi.org/10.1109/TUFFC.2008.819.
[32] M. Muller, P. Moilanen, E. Bossy, et al., Comparison of three ultrasonic axial transmission methods for bone assessment, Ultrasound Med. Biol. 31 (2005) 633–642, https://doi.org/10.1016/j.ultrasmedbio.2005.02.001.
[33] P. Nauleau, J.-G. Minonzio, M. Chekroun, et al., A method for the measurement of dispersion curves of circumferential guided waves radiating from curved shells: experimental validation and application to a femoral neck mimicking phantom, Phys. Med. Biol. 61 (2016) 4746–4762, https://doi.org/10.1088/0031-9155/61/13/4746.
[34] M. Chekroun, J.-G. Minonzio, C. Prada, et al., Measurement of dispersion curves of circumferential guides waves radiation from curved shells: theory and numerical validation, J. Acoust. Soc. Am. (2016) doi: 10.1121/1.4941652.
[35] P. Moilanen, P.H.F. Nicholson, V. Kilappa, et al., Assessment of the cortical bone thickness using ultrasonic guided waves: modelling and in vitro study, Ultrasound Med. Biol. 33 (2007) 254–262, https://doi.org/10.1016/j.ultrasmedbio.2006.07.038.
[36] E. Bossy, M. Talmant, M. Defontaine, et al., Bidirectional axial transmission can improve accuracy and precision of ultrasonic velocity measurement in cortical bone: a validation on test materials, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 51 (2004) 71–79, https://doi.org/10.1109/TUFFC.2004.1268469.
[37] A. Tatarinov, V. Egorov, N. Sarvazyan, A. Sarvazyan, Multi-frequency axial transmission bone ultrasonometer, Ultrasonics (2013), https://doi.org/10.1016/j.ultras.2013.09.025.
[38] H. Tuna, M. Birtane, G. Ekuklu, et al., Does quantitative tibial ultrasound predict low bone mineral density defined by dual energy X-ray absorptiometry? Yonsei Med. J. 49 (2008) 436–442, https://doi.org/10.3349/ymj.2008.49.3.436.
[39] C.R. Krestan, S. Grampp, C. Henk, et al., Limited diagnostic agreement of quantitative sonography of the radius and phalanges with dual-energy X-ray absorptiometry of the spine, femur, and radius for diagnosis of osteoporosis, Am. J. Roentgenol. 183 (2004) 639–644, https://doi.org/10.2214/ajr.183.3.1830639.
[40] K.M. Knapp, G.M. Blake, T.D. Spector, I. Fogelman, Can the WHO definition of osteoporosis be applied to multi-site axial transmission quantitative ultrasound? Osteoporos. Int. 15 (2004) 367–374, https://doi.org/10.1007/s00198-003-1555-4.
[41] R. Barkmann, E. Kantorovich, C. Singal, et al., A new method for quantitative ultrasound measurements at multiple skeletal sites: first results of precision and fracture discrimination, J. Clin. Densitom. 3 (2000) 1–7.
[42] D. Hans, S.K. Srivastav, C. Singal, et al., Does combining the results from multiple bone sites measured by a new quantitative ultrasound device improve discrimination of hip fracture? J. Bone Miner. Res. 14 (1999) 644–651, https://doi.org/10.1359/jbmr.1999.14.4.644.
[43] M. Talmant, S. Kolta, C. Roux, et al., In vivo performance evaluation of bi-directional ultrasonic axial transmission for cortical bone assessment, Ultrasound Med. Biol. 35 (2009) 912–919, https://doi.org/10.1016/j.ultrasmedbio.2008.12.008.
[44] F. Vogl, B. Bernet, D. Bolognesi, W.R. Taylor, Towards assessing cortical bone porosity using low-frequency quantitative acoustics: a phantom-based study, PLoS One 12 (2017).
[45] E. Stüssi, D. Fäh, Assessment of bone mineral content by in vivo measurements of flexural wave velocities, Med. Biol. Eng. Compu. 26 (1988) 349.
[46] N.M. Keshawarz, R.R. Recker, Expansion of the medullary cavity at the expense of cortex in postmenopausal osteoporosis, Metab Bone Dis Relat Res 5 (1984) 223–228, https://doi.org/10.1016/0221-8747(84)90063-8.
[47] H. Ritzel, M. Amling, M. Pösl, et al., The thickness of human vertebral cortical bone and its changes in aging and osteoporosis: a histomorphometric analysis of the complete spinal column from thirty-seven autopsy specimens, J. Bone Miner. Res. 12 (1997) 89–95, https://doi.org/10.1359/jbmr.1997.12.1.89.
[48] Y. Bala, R. Zebaze, A. Ghasem-Zadeh1, et al., Cortical porosity identifies women with osteopenia at increased risk for forearm fractures, J. Bone Min. Res. 29 (2014) 1356–1362, https://doi.org/10.1038/jid.2014.371.
[49] D.M.L. Cooper, C.E. Kawalilak, K. Harrison, et al., Cortical bone porosity : what is it, why is it important, and how can we detect it ? Curr. Osteoporos. Rep. 187–198 (2016), https://doi.org/10.1007/s11914-016-0319-y.
[50] K. Raum, I. Leguerney, F. Chandelier, et al., Bone microstructure and elastic tissue properties are reflected in QUS axial transmission measurements, Ultrasound Med. Biol. 31 (2005) 1225–1235, https://doi.org/10.1016/j.ultrasmedbio.2005.05.002.
[51] C.F. Njeh, I. Saeed, M. Grigorian, et al., Assessment of bone status using speed of sound at multiple anatomical sites, Ultrasound Med. Biol. 27 (2001) 1337–1345.
[52] A. Sarvazyan, A. Tatarinov, V. Egorov, et al., Application of the dual-frequency ultrasonometer for osteoporosis detection, Ultrasonics 49 (2009) 331–337, https://doi.org/10.1016/j.ultras.2008.10.003.
[53] C.M. Bishop, Pattern Recognition and Machine Learning, 1st ed., Springer, Dordrecht Heidelberg London New York, 2013.
[54] A. Singh, M.K. Dutta, R. Jennane, E. Lespessailles, Classification of the trabecular bone structure of osteoporotic patients using machine vision, Comput. Biol. Med. 91 (2017) 148–158, https://doi.org/10.1016/j.compbiomed.2017.10.011.
[55] C. Bahlmann, B. Haasdonk, H. Burkhardt, Online handwriting recognition with support vector machines – a kernel approach, in: Proc. Eighth Int. Work. Front. Handwrit. Recognit., 2002, pp 49–54.
[56] K. Virupakshappa, E. Oruklu, Ultrasonic flaw detection using Support Vector Machine classification, in: 2015 IEEE Int. Ultrason. Symp., 2015, pp 1–4.
[57] M. Denis, L. Wan, M. Fatemi, A. Alizad, Ultrasound characterization of bone demineralization using a support vector machine, Ultrasound Med. Biol. 44 (2018) 714–725, https://doi.org/10.1016/j.ultrasmedbio.2017.11.004.
[58] F. Vogl, K. Schnüriger, H. Gerber, W.R. Taylor, Reliability of phase velocity measurements of flexural acoustic waves in the human tibia in-vivo, PLoS One 11 (2016) e0152417, https://doi.org/10.1371/journal.pone.0152417.
[59] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[60] S.H. Prins, H.L. Jørgensen, L.V. Jørgensen, C. Hassager, The role of quantitative ultrasound in the assessment of bone: a review, Clin. Physiol. 18 (1998) 3–17.
[61] T.D. Szwedowski, W.R. Taylor, M.O. Heller, et al., Generic rules of mechano-regulation combined with subject specific loading conditions can explain bone adaptation after THA, PLoS One (2012), https://doi.org/10.1371/journal.pone.0036231.
[62] R. Barkmann, P. Laugier, U. Moser, et al., A method for the estimation of femoral bone density from quantitative ultrasound variables measured directly at the human femur, Bone 40 (2007) 37–44.
[63] A.K. Amstrup, N.F.B. Jakobsen, E. Moser, et al., Association between bone indices assessed by DXA, HR-pQCT and QCT scans in post-menopausal women, J. Bone Miner. Metab. 34 (2016) 638–645, https://doi.org/10.1007/s00774-015-0708-9.
[64] X.S. Liu, A. Cohen, E. Shane, et al., Bone density, geometry, microstructure and stiffness: relationships between peripheral and central skeletal sites assessed by DXA, HR-pQCT, cQCT in premenopausal women, J. Bone Miner. Res. 25 (2010) 2229–2238, https://doi.org/10.1002/jbmr.111.Bone.
[65] R.B. Cook, D. Collins, J. Tucker, P. Zioupos, The ability of peripheral quantitative ultrasound to identify patients with low bone mineral density in the hip or spine, Ultrasound Med. Biol. 31 (2005) 625–632, https://doi.org/10.1016/j.ultrasmedbio.2005.02.003.
[66] S. Nayak, I. Olkin, H. Liu, et al., Meta-analysis: accuracy of quantitative ultrasound for identifying patients with osteoporosis, Ann. Intern. Med. 11 (2006) 832–841.
[67] O. Ilic Stojanovic, M. Vuceljic, M. Lazovic, et al., Bone mineral density at different sites and vertebral fractures in Serbian postmenopausal women, Climacteric (2016) 1–7, https://doi.org/10.1080/13697137.2016.1253054.
[68] K.M. Knapp, G.M. Blake, T.D. Spector, I. Fogelman, Multisite quantitative ultrasound: precision, age- and menopause-related changes, fracture discrimination, and T-score equivalence with dual-energy X-ray absorptiometry, Osteoporos. Int. 12 (2001) 456–464, https://doi.org/10.1007/s001980170090.
[69] M.E. Tipping, Sparse Bayesian learning and the relevance vector machine, J. Mach. Learn. Res. 1 (2001) 211–244, https://doi.org/10.1162/15324430152748236.