Quantitative ultrasound for the detection and management of osteoporosis

Didier Hans, PhD, PD,(1) Marc-Antoine Krieg, MD, PD.(1)

Abstract
Quantitative ultrasound (QUS) appears to be developing into an acceptable, low-cost and readily-accessible alternative to dual X-ray absorptiometry (DXA) measurements of bone mineral density (BMD) in the detection and management of osteoporosis. Perhaps the major difficulty with their widespread use is that many different QUS devices exist that differ substantially from each other, in terms of the parameters they measure and the strength of empirical evidence supporting their use. But another problem is that virtually no data exist outside of Caucasian or Asian populations. In general, heel QUS appears to be most tested and most effective. Some, but not all heel QUS devices are effective assessing fracture risk in some, but not all populations, the evidence being strongest for Caucasian females > 55 years old, though some evidence exists for Asian females > 55 and for Caucasian and Asian males > 70. Certain devices may allow to estimate the likelihood of osteoporosis, but very limited evidence exists supporting QUS use during the initiation or monitoring of osteoporosis treatment. Likely, QUS is most effective when combined with an assessment of clinical risk factors (CRF); with DXA reserved for individuals who are not identified as either high or low risk using QUS and CRF. However, monitoring and maintenance of test and instrument accuracy, monitoring and maintenance of test and instrument accuracy,

Resumen
El ultrasonido cuantitativo (QUS) es una alternativa para la detección y manejo de la osteoporosis de bajo costo y uso práctico, si se compara con las densitometrías de rayos X de doble haz de baja energía (DXA) que determinan densidad mineral ósea (BMD). La mayor dificultad para el uso generalizado del QUS por un lado es que existen muchos instrumentos que son significativamente diferentes uno del otro y por otro en la calidad de la evidencia en que se justifica su empleo, que generalmente es insuficiente y/o poco sistematizada. Otro problema importante del QUS, es que prácticamente no existe información que no sea la generada en poblaciones asiáticas o caucásicas. En general, los estudios de calcáneo realizados con QUS son los más utilizados y mejor validados para evaluar el riesgo de fracturas en algunas poblaciones. La evidencia más grande de su efectividad se conoce para las mujeres caucásicas y asiáticas mayores de 55 años e incluso para los hombres asiáticos mayores de 70 años. Varios instrumentos cuentan con buen sustento científico, que los vuelven confiables para establecer un pronóstico preciso e identificar el riesgo individual de sufrir fracturas por osteoporosis, sin embargo, existe poca evidencia que respalde su uso para iniciar y monitorear el resultado del tratamiento de la osteoporosis. El QUS mejora su efectividad diagnóstica cuando se combina

Part of this work was written in preparation of the 2007 Position Conference Development (chaired by Sandy Baim) of the International Society Clinical Densitometry with the participation of Reinhart Barkmann, Stefano Gonnelli, Alison Stewart, Douglas C. Bauer, Luis Del Rio Barquero, Jonathan J. Kaufman, E. Michael Lewiecki, Roman Lorenz, Paul D. Miller, Wojciech P. Olszynski, Catalina Poiana, Anne-Marie Schott.

It does not represent the official position statements of ISCD on the use of QUS in clinical routine. Official position statements issued from the ISCD 2007 Position Conference Development were published in the Journal of Clinical Densitometry early 2008.

(1) Center of Bone Diseases, Lausanne University Hospital, Lausanne, Switzerland.
Osteoporosis is a “disease characterized by low bone mass and micro-architectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk”. Hip fractures are especially problematic. In the US, for example, more than 250,000 hip fractures occur annually; at least 90% of which are attributed to osteoporosis; women over 70 years of age are particularly vulnerable.5 Only 15% of patients are able to walk without assistance 6 months after the event; 50% never return to their previous functional state; and roughly 20% require long-term care. Hip fracture patients also experience a 20% increase in mortality over the next five years. Even in relatively smaller-population countries like Canada, costs to governments measure in the billions of dollars annually, related both to direct health care costs and insurance. And, as much of a problem hip fractures currently are, the numbers almost certainly will increase dramatically as an increasing percentage of the population achieves older age, with as many as 6.3 million hip fractures predicted worldwide, annually, by 2050.16

Osteoporosis is a major public health concern in Latin America as well, with vertebral osteoporosis affecting 12-18% and femoral osteoporosis 8-22% of women 50 years and older. Bone mineral density may be lower in Latin American women over 50 than in their American counter-parts, with osteopenia affecting almost 60% of women ≥ 50. In addition, in Latin America, up to 362 osteoporosis-related hip fractures occur annually per 100,000 persons 50 years and older. Vertebral fractures affect almost one in five women over 50, and between 17 and 37% of hip fracture suffers die within a year of their fracture. The social burden of osteoporosis also is high in South and Central America.
with DXA, which include difficulties extrapolating standards for hip fracture risk to other skeletal sites, like the wrist and lumbar spine. Two other major problems with DXA that are especially pertinent in South and Central America are (1) its cost, and (2) the rarity of DXA instruments in many localities, especially in poorer and/or rural areas and in less developed countries. These two latter problems have led many investigators to search for some lower-cost and more readily-available alternative to DXA for the diagnosis of osteoporosis and/or the estimation of future risk of fragility fractures; and one such alternative that has garnered considerable recent attention has been quantitative ultrasound (QUS).

Quantitative ultrasound: General principles

Initially used to detect enemy submarines underwater during World War II, ultrasonic waves are sound waves outside the threshold of human hearing, which pass easily through fluid and other tissues, and which are altered upon contact with bone, in terms of their shape, intensity and speed. Over the years, ultrasonic (US) devices have found a diverse array of clinical applications in medicine, including uses in cardiology to assess cardiac size and function and vascular flow, obstetrics to assess fetal development, general medicine to examine for intra-abdominal and intra-peritoneal masses, and rheumatology and orthopedics to both diagnose and treat conditions like bursitis and tendinitis. As opposed to qualitative ultrasound, which just generates pictures, quantitative ultrasound uses ultrasonic waves at lower frequencies to generate empirical measurements.

With respect to the detection of osteoporosis, QUS can be used to measure a variety of parameters that pertain to bone density, parameters that are related to the velocity and attenuation of US waves as they pass through bone. Advantages of QUS over DXA are that it is inexpensive, transportable, and ionizing radiation free. The low cost and transportability could make QUS an especially valuable osteoporosis detection tool wherever cost or instrument inaccessibility renders DXA difficult or impossible. But does QUS work? Already, there is evidence that QUS is as effective as axial DXA in predicting hip fractures and all osteoporosis-related fractures in elderly women.24-26 Having said this, numerous potential problems still exist with the use of QUS for osteoporosis detection. For example, care must be exercised interpreting US velocity and attenuation, as they are calculated differently depending upon the manufacturer and model of the ultrasound device. Similarly, there are significant differences between QUS instruments from different manufacturers, differences that affect the interpretation of results and limit comparisons between devices.

Different QUS devices

QUS devices can be classified into three types, related to the form of US transmission used:

1. Trabecular sound transmission is the most commonly utilized category of devices, for which the most evidence exists supporting its use. It is best utilized measuring the heel.24,25
2. Cortical transverse transmission currently only is used in phalax contact devices;27 to date, little evidence supports the use of these devices clinically for osteoporosis.
3. Cortical axial transmission presently is being investigated for use in phalanges, the radius and the tibia;27 no clinical application have been proven to date.

As just noted, heel devices currently appear to have the most clinical applications, with some devices –like the GE-Lunar Achilles and the Hologic Sahara– better tested and more proven effective than others (table 1). For these purposes, the recommended parameter of interest generally has been the heel stiffness index (SI) or the Quantitative index (QUI), which is a composite score combining the results of broadband ultrasound attenuation (BUA) and speed of sound (SOS), as measured in meters per second.

The remainder of this paper will review the clinical use of QUS in the following settings: 1) the prediction of fracture risk; 2) the diagnosis of osteoporosis; 3) the initiation of osteoporosis treatment or prevention; 4) the monitoring of such treatment; and 5) osteoporosis case finding. The paper will conclude by examining 6) quality assurance and quality control issues pertaining to the clinical application of QUS.

1) Using QUS to predict fracture risk

At the present time, there is satisfactory (cross-sectional and/or prospective) evidence that QUS can be used to assess fracture risk in some, but not all populations, as defined by sex, age and ethnic background.26-96 This is particularly true of heel QUS and for hip versus spinal fractures. Having said this, because of various methodological issues, it is difficult to compare studies. Nonetheless, combining the results from 13 studies involving 9 561 patients,55,39,40,43,45,48,52,58-60,69,79 it is reasonable to state that the increase in relative risk observed for each standard deviation decrease in stiffness index (SI),
measured at the heel using QUS, is roughly 2.0 for the hip and spine, and approximately 1.5 for all fractures combined. Consequently, heel QUS is much the same as DXA BMD, in terms of hip and spine fracture risk per standard deviation decrease.\textsuperscript{97,98}

Although some differences may exist in the expression of osteoporosis and overall fracture risk in Hispanics versus general Caucasians and other ethnic populations,\textsuperscript{99,100} there is ample empirical evidence that the heel QUS stiffness index, using some but not all QUS devices, is predictive of hip fracture risk in Caucasian and Asian women over age 55, and of any fracture risk in Caucasian women > 55 (table II). Weaker evidence exists that the heel QUS stiffness index, again using some but not all QUS devices, is predictive of hip fracture risk in Caucasian and Asian men over age 70; of vertebral fracture risk in Caucasian and Asian women over age 55; and of any fracture risk in Asian women and Caucasian or Asian men > 70. With respect to QUS devices from one of the other two categories, phalanx QUS devices utilizing cortical transverse transmission might predict non-vertebral fracture risk in Caucasian women > 70; however, to date, cortical axial transmission devices have no proven clinical application.

Another practical question is: which QUS device is best to use? As indicated in table I, the GE Lunar Achilles and the Hologic Sahara are among the best tested devices, at the hip, spine and overall, and both seem effective for most females; the former may be preferable in males. Some evidence exists supporting the use of the Norland Cuba Clinical and the IGEA DBM Sonic BP, at least among Caucasians. However general results on the IGEA DBM Sonic device are not very impressive. Consequently, the three heel devices appear to be the most reasonable to use, at this time, though further testing of these and other devices clearly is necessary.

2) Using QUS to diagnose osteoporosis

Diagnosing osteoporosis using QUS is less supported by evidence and more complicated and problematic than assessing fracture risk is. To start with, the T-score diagnostic criteria of -2.5, classically used for DXA BMD, cannot be applied to QUS without discrepancies in the numbers of women diagnosed with osteoporosis. This is because there are tremendous variations in QUS measurements by skeletal site, and because different QUS devices yield different results. If, for example, the prevalence of osteoporosis is defined as -2.5 standard deviations from the mean threshold for QUS, even within the same sample population, different QUS instruments and different skeletal sites generate prevalence estimates that vary as much as ten-fold; for example, prevalence estimates among Caucasian women > 65 have ranged

---

**Table I**

**Quantitative Ultrasound (QUS) Devices Currently Available**

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Model</th>
<th>Ability to assess hip fracture risk</th>
<th>Ability to assess spine fracture risk</th>
<th>Ability to assess all fracture risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE-medical (Lunar)</td>
<td>Achilles</td>
<td>Proven in most populations</td>
<td>Proven in most populations</td>
<td>Proven in most populations</td>
</tr>
<tr>
<td>DMS</td>
<td>Ubis 3000/5000</td>
<td>Some evidence</td>
<td>Some evidence</td>
<td>Some evidence</td>
</tr>
<tr>
<td>Hologic (McCue)</td>
<td>Sahara</td>
<td>Proven in Caucasians</td>
<td>Proven in Caucasian females</td>
<td>Proven in Caucasians</td>
</tr>
<tr>
<td>Norland (McCue)</td>
<td>Cuba Clinical</td>
<td>Proven in Caucasians</td>
<td>Some evidence</td>
<td>Proven in Caucasians</td>
</tr>
<tr>
<td>IGEA</td>
<td>DBM Sonic BP</td>
<td>Proven in Caucasian females</td>
<td>Proven in Caucasian females</td>
<td>Some evidence</td>
</tr>
<tr>
<td>BeamMed (Sunlight)</td>
<td>Omnisense</td>
<td>Some evidence</td>
<td>Some evidence</td>
<td>Some evidence</td>
</tr>
<tr>
<td>Meditech</td>
<td>DTU-One</td>
<td>No evidence</td>
<td>Some evidence</td>
<td>Some evidence</td>
</tr>
<tr>
<td>Aloka</td>
<td>AOS-100</td>
<td>Some evidence</td>
<td>No evidence</td>
<td>No evidence</td>
</tr>
<tr>
<td>Medilink</td>
<td>Osteospace</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quidel Inc.</td>
<td>QUS-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ishikawa Seisakusho</td>
<td>CM-100 / 200</td>
<td>No Evidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elk Co.</td>
<td>Sonost 2000 / 3000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Osteosys Co.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPtech21 Co.</td>
<td>Osteolmager Plus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPtech21 Co.</td>
<td>Osteo Pro</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data gleaned from Quantitative Ultrasound (QUS) in the Management of Osteoporosis: The 2007 ISCD Official Positions
from 4 to 50%.[96,101-105] To overcome this dilemma, there is a need for pre-defined, device-specific diagnostic thresholds. One recommended system suggests calibrating QUS measurements with DXA results, the latter used as the 'gold standard', so that an upper QUS threshold is set to identify osteoporosis with 90% sensitivity, and a lower threshold is set to identify osteoporosis with 90% specificity.[106,107] A similar approach already has been recommended by the UK National Osteoporosis Society to define upper and lower thresholds for pDXA, the results of which are highly correlated with QUS. Using such a system, one could identify osteoporosis with high probability in patients whose results fall below the lower threshold for QUS, where specificity exceeds 90%; between the upper and lower thresholds, the diagnosis of osteoporosis would be considered quite equivocal, so that another means of measurement, like DXA BMD, would be highly recommended; and above the upper threshold for QUS, where the sensitivity of a value below the threshold is 90%, osteoporosis would be deemed unlikely.

We, in fact, utilized this approach in 5,954 women 75 years and older who took part in the EPIDOS Study, utilizing the 90% sensitivity threshold for the Achilles stiffness index of SI = 78%, and the 90% specificity threshold of SI = 57%. Using these cut-off points generated 11% false positive (FP) and 13% false negative (FN) results, which are comparable to the FP and FN rates of many other tests. Based upon these results, we believe that device-specific heel QUS thresholds for 90% sensitivity and 90% specificity in specific populations defined by sex, age, and ethnic background, can be used to identify individuals who have either a high or a low likelihood of osteoporosis, even though only limited evidence exists supporting the use of any existing QUS device for this purpose. Devices that have been evaluated include the GE-Lunar Achilles, the Hologic Sahara, and the DMS UBIS-5000, each of which uses a different measure and different upper and lower likelihood thresholds, as indicated in table III. What can be concluded is that, regardless of the QUS device used, values that fall between the upper and lower thresholds strongly warrant further evaluation using DXA BMD as a more definitive test.

3) Using QUS to initiate osteoporosis treatment

Except in patients with a low-energy fracture of the hip or spine, when the fracture alone is adequate to require treatment, all currently-published guidelines or recommendations for the initiation of osteoporosis treatment are based upon DXA BMD values; in no instance, to date, are the results of QUS the definitive parameter. Despite this, several studies have demonstrated high levels of correlation (r ~ 0.90) between heel trabecular sound transmission and BMD at matched skeletal sites.[110-113] Moreover, both SOS and BUA, standard QUS measurements, are dependent upon overall bone strength which, in turn, is related to bone density, architecture and turnover, and the extent of bone mineralization.[110,112,114-121] These factors likely work together to maintain overall bone quality and strength, and to prevent fractures and other bone failure. [110,112,114-121] QUS parameters related to heel trabecular transverse transmission are highly correlated with bone strength.[117,122-129] Consequently, it is conceivable that guidelines could be created using QUS to guide when to initiate osteoporosis treatment,

### Table II
**Using Quantitative Ultrasound (QUS) to Assess Fracture Risk**

<table>
<thead>
<tr>
<th>Device</th>
<th>Population studied</th>
<th>Skeletal site studied</th>
<th>Strength of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>CF &gt; 55 y/o</td>
<td>Hip</td>
<td>Good</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>AF &gt; 55 y/o</td>
<td>Hip</td>
<td>Fair</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>CM &gt; 70 y/o</td>
<td>Hip</td>
<td>Fair</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>AM &gt; 70 y/o</td>
<td>Hip</td>
<td>Fair</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>CF &gt; 55 y/o</td>
<td>Spine</td>
<td>Good</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>AF &gt; 55 y/o</td>
<td>Spine</td>
<td>Poor</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>CM &gt; 70 y/o</td>
<td>Overall</td>
<td>Good</td>
</tr>
<tr>
<td>Some, but not all heel QUS devices</td>
<td>AM &gt; 70 y/o</td>
<td>Overall</td>
<td>Good</td>
</tr>
<tr>
<td>Phalanx cortical transverse transmission devices</td>
<td>CF &gt; 70 y/o</td>
<td>Non-vertebral</td>
<td>Fair</td>
</tr>
</tbody>
</table>

CF= Caucasian females; AF= Asian females; CM= Caucasian males; AM= Asian males

### Table III
**Using Quantitative Ultrasound (QUS) to Diagnose Osteoporosis**

<table>
<thead>
<tr>
<th>Device</th>
<th>Population studied</th>
<th>Threshold for high likelihood of osteoporosis</th>
<th>Threshold for low likelihood of osteoporosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE-Lunar Achilles</td>
<td>CF and AF &gt; 65</td>
<td>SI ≤ 57.0</td>
<td>SI &gt; 78</td>
</tr>
<tr>
<td>Hologic Sahara</td>
<td>CF and AF &gt; 65</td>
<td>QUI ≤ 59.0</td>
<td>QUI &gt; 83</td>
</tr>
<tr>
<td>DMS UBIS-5000</td>
<td>CF and AF &gt; 65</td>
<td>BUA ≤ 55.0</td>
<td>BUA &gt; 62</td>
</tr>
</tbody>
</table>

CF= Caucasian females; AF= Asian females
especially if combined with the use of clinical risk factors.\textsuperscript{130} To date, however, no randomized clinical trails have been published examining whether individuals identified as high risk for fracture by QUS respond to treatment.

At the present time, published evidence does not support using QUS device-specific values to initiate osteoporosis treatment in women younger than 65 years or in men of any age, but some evidence exists supporting QUS use in other populations, when DXA is not available. These populations include 1) Caucasian and 2) Asian women who are between the ages of 65 and 74 years, exhibit results below the lower specific device threshold (i.e. SI ≤ 57.0 with the GE-Lunar Achilles device), and have at least two clinical risk factors (table IV); and 3) Caucasian and 4) Asian women who are 75 years or older, exhibit results below the lower specific device threshold (i.e. SI ≤ 57.0 with the GE-Lunar Achilles device), and have at least one clinical risk factor besides age. Table IV lists pertinent clinical risk factors that we have identified by examining published meta-analyses, as well as literature reviews written by Kanis\textsuperscript{131} and Durosier.\textsuperscript{98}

Currently, a World Health Organization (WHO) task force is developing a model to predict the 10-year probability of osteoporosis-related fractures, combining femoral neck DXA BMD measurements and CRF. This also could be done combining QUS and CRF, using a device-specific T-score. QUS-generated high and low risk probabilities – for example, SI ≤ 57.0 and > 78, respectively, for the GE-Lunar Achilles device– then could be used to decide whether treatment is warranted. Because different devices have their own high-and low-probability threshold values that all correspond to roughly the same two levels of fracture risk, this approach could be utilized independently of the measurement instrument used. Preliminary results demonstrating the benefits of this combined CRF plus technology approach already have been published by Hans \textit{et al.}\textsuperscript{142}

Unfortunately, whereas QUS and BMD are highly correlated in trabecular bone\textsuperscript{143} and this correlation reasonably well understood, the situation is considerably more complex with cortical measurements. Many properties influence these measurements, including cortical thickness, mineralization, porosity and lamellar structure, and it is not clear to what degree these various properties contribute to bone strength.\textsuperscript{144-147} Consequently, QUS devices that measure cortical bone, like the cortical transverse transmission devices currently used to assess phalanges, and the cortical axial transmission devices being investigated for use with phalanges, the radius and tibia, cannot be recommended as tools to determine the appropriateness of initiating osteoporosis treatment at this time.

4) Using QUS to monitor osteoporosis treatment

At this time, QUS cannot be recommended for the monitoring of treatment response in patients with osteoporosis, both due to the absence of large-scale, randomized, double-blinded and placebo-controlled clinical trials (RCT) and the relatively equivocal evidence that has been generated by the studies that have been published.\textsuperscript{148-164} It has been observed that changes in heel QUS parameters, especially the stiffness index (SI), do mimic the treatment response observed in BMD. In two studies involving alendronate, for example, the Achilles SI was observed to significantly increase with treatment over time.\textsuperscript{149,150} Clearly, however, further RCT are needed to determine if QUS parameters are sensitive enough to change with treatment, if the various QUS instruments are sensitive enough to detect these changes, and if the precision of these instruments is such that repeated measures can be performed without excessive ‘noise’. There is some evidence that instrument precision is adequate in the short-term; but what about over a longer period of time? What can be said is that, if QUS is going to be used to monitor treatment, likely the heel devices will be most successful, since trabecular measurements appear to be more accurate than those that have been achieved with any of the cortical devices.

5) Case finding

Case finding involves distinguishing subjects at highest or lowest risk for a given disorder, who hence do not require further investigation because their disease

---

**Table IV**

**Clinical risk factors for osteoporosis for use with Quantitative Ultrasound**

- Age over 75 years\textsuperscript{132,133}
- Low BMI (<20 kg/m\textsuperscript{2})\textsuperscript{132,134,135}
- Previous fragility fracture after the age of 50\textsuperscript{132,135,136}
- Maternal history of hip fracture\textsuperscript{137}
- Current smoking\textsuperscript{138}
- Diabetes
- No prior hormone replacement therapy (HRT)\textsuperscript{135}
- Prior use of systemic glucocorticoids\textsuperscript{135,139}
- Fall within the past 12 months\textsuperscript{135,140}
- Use of arms to stand up 3 times from a chair (Missed Chair Test)\textsuperscript{135,140,141}
status is reasonably well known, from subjects at intermediate risk whose disease status remains equivocal, thereby requiring further evaluation. At present, there is fair empirical support for the use of heel QUS for osteoporosis case finding among Caucasian and Asian females who are at least 65 years old. The evidence for the use of cortical devices is poor, and for males and other ethnic populations generally lacking altogether. Nonetheless, we feel that the proposed case-finding protocol depicted in figure 1 is reasonable for clinical practice, in terms of distinguishing individuals who require from those who do not require further evaluation of fragility fracture risk.

To begin, each patient would undergo an assessment to identify any clinical risk factors. Using this protocol, patients who are suspected to be at risk for secondary osteoporosis—for example, because of prolonged systemic use of corticosteroids—or who have a clinically-evident vertebral fracture would proceed directly to management deemed appropriate for their condition. All others would undergo a heel QUS. Based upon this, those whose QUS parameters suggest a low likelihood of future fracture would be assigned to receive primary prevention, unless they have had a fragility fracture, in which case they would undergo DXA BMD. Those for whom results indicate an intermediate risk of future
fracture, in that they lie between the upper and lower thresholds, also would undergo DXA BMD. And those patients for whom the risk of future fracture is deemed to be high, based upon the results of heel QUS, would proceed to treatment if they are ≥ 75 years old and have at least one other CRF besides age, or if they are < 75 and have at least two clinical risk factors. Otherwise, they too would proceed to DXA BMD.

The primary advantage of this protocol, especially in terms of cost, is that it saves performing costly and, sometimes relatively inaccessible DXA on all patients. In particular, those with a low risk of fracture by QUS would avoid DXA unless they have had a fragility fracture; and those with a high risk of fracture by QUS would avoid DXA if they have at least two CRF, counting age. As stated at the outset, this would have particular relevance in localities in which access to DXA is scarce or too costly to be performed, both of which likely apply to the majority of patients in many developing countries.

6) QUS quality assurance and quality control

Technically, quality assurance primarily deals with the performance of the equipment, whereas quality control is more heavily grounded in theory and statistics, emphasizing the quality of the actual test. For practical purposes, however, these two concepts often are treated as the same, and this paper will not seek to further delineate them. The primary issues of importance are those of test/equipment accuracy, precision, and reproducibility. Accuracy is a measure of how close a provided answer or value is to the true answer or value. If one were to visualize a game of darts, for example, darts that hit the bull’s eye are said to be accurate. Precision, on the other hand, is a measure of consistency. If five darts are thrown, do they all end up close to each other (high precision) or widely spread (low precision)? The two concepts, accuracy and precision, are not the same. One could have all one’s darts bunched together far from the bull’s eye (high precision, low accuracy); or one could have all one’s darts widely spread, but all equidistant from the bull’s eye, so that the average of their positions is near to the center dot (high accuracy, low precision). Reproducibility measures how well the same test done on the same person or sample yields the same result, whether the test is performed by the same technician (intra-observer) or a different technician (inter-observer). These three parameters are important, whether a test is being performed to diagnose disease, monitor its course, or identify potential cases.

With heel QUS, there are several potential sources of in vivo measurement error, which include surrounding soft tissue and foot positioning, soft-tissue thickness, temperature, and composition; the quality of sound transmission from the coupling medium to the skin; and properties of the coupling medium between the transducers and the skin, whether it be a fluid bath or sound transmitting pads. One of the most important components of QA entails using some sort of test object, which can be either a standard or phantom, to monitor instrument performance and make necessary calibrations when measurement accuracy begins to stray. A standard is an object of known acoustic properties, which does not necessarily resemble the anatomy of interest. Conversely, a phantom is designed to emulate the anatomy and acoustic properties that exist during in vivo measurements as much as possible. With respect to optimizing QUS device performance, phantoms are more useful. Unfortunately, no universally-applicable phantoms exist. Consequently, manufacturer-specific phantoms must be used, and measured each day that the respective device is used, following the manufacturer’s protocol, to detect performance changes that may result from component aging or failure. Detection of these changes allows both for necessary repairs, and for adjustments to specific readings by applying a correction factor to patient data. Table V provides a brief list of important guidelines regarding quality assurance and control:

**Summary**

To date, no satisfactory evidence exists either supporting or refuting the usefulness of quantitative ultrasound in Latin populations, so that further research clearly is warranted. Nonetheless, there is enough evidence in other populations to suggest that QUS may be an acceptable, low-cost and readily-accessible alternative to DXA measurements of BMD in the management of osteoporosis in Hispanics. Many different QUS devices exist that are quite different in terms of the parameters they measure and the strength of empirical evidence supporting their use. In general, heel QUS appears to be most tested and most effective. Some, but not all heel QUS devices are effective assessing fracture risk in some, but not all populations, the evidence being strongest for Caucasian females > 55 years old, though some evidence exists for Asian females > 55 and for Caucasian and Asian males > 70. Certain devices may allow for the accurate diagnosis of osteoporosis, but very limited evidence exists supporting the use of QUS use during the initiation or monitoring of osteoporosis treatment. A reasonable protocol for osteoporosis case-finding relies upon the combined assessment of clinical risk factors and heel QUS. However, monitoring and maintenance of test and
Quantitative ultrasound and osteoporosis

Table V

RECOMMENDATIONS FOR QUALITY ASSURANCE/QUALITY CONTROL IN QUS SYSTEMS

(1) QUS-specific information and training must be provided prior to measuring patients in routine clinical practice;
(2) To assist in optimizing QUS measurements, in vivo precision error should be calculated by each QUS facility’s technologist; in accordance with ISCD recommendations for DXA, rather than relying upon precision values reported in the literature or by the manufacturer; once measurement precision is known, it is possible to determine the minimal detectable difference between two measurements that is statistically significant;
(3) Obtaining facility accreditation for QUS is recommended;
(4) As with other methods, with QUS, anatomically-consistent regions of the bone must be measured to enhance accuracy; and precision is affected by the quality of sequential repositioning of each patient; the use of an imaging system may help to overcome positioning errors by placing a region of interest (ROI) on the image;79
(5) Attempts should be made to maintain a relatively-constant and standard temperature for all readings; and when an alteration is felt to exist because of temperature, accommodation should be made and the measurement repeated; for example, with heel measurements, a cold foot can falsely lower SOS measurements; hence, if the SOS result is low and this value clinically-important, the foot should be warmed and the measurement repeated;
(6) Oedema of the foot can significantly affect heel QUS readings, so that this test should not be performed in such patients;
(7) Achieving good-quality coupling of the ultrasonic beam as it passes into the body is essential; a coupling gel or liquid must be used with all methods to ensure proper penetration through the skin;
(8) Manufacturer-specific phantoms must be measured each day a device is used, following the manufacturer’s protocol, to detect equipment changes that may result from component aging or failure, and thereby allow for necessary repairs, or for adjustments to specific readings by applying a correction factor to patient data;177
(9) It generally is accepted that long-term stability and precision errors detected during phantom measurement should not exceed ¼ T-score to restrict misclassification errors to within a clinically-acceptable level119
(10) Each facility should periodically assess the performance of its technologists; two essential components of this are (1) ensuring that adequate time is utilized performing measurements; and (2) measuring intra-observer reliability to generate a coefficient of variation (CV); the in vivo CVs generally quoted in the literature range from 0.8 - 5 % and 0.2 - 1% for BUA and SOS, respectively.

References


130. Durosiere C, et al. Combining clinical factors and QUS improves the detection of women both at low and high risk for hip fracture. [submitted].


