Osteoporos Int (2005) 16: 435–445
DOI 10.1007/s00198-004-1683-5

ORIGINAL ARTICLE

Appropriateness of internal digital phantoms for monitoring the stability of the UBIS 5000 quantitative ultrasound device in clinical trials

D. Hans · I. Alekxandrova · C. Njeh · C. Le-Floch
G. Conicella · P. Laugier · D.O. Slosman

Received: 10 December 2003 / Accepted: 19 May 2004 / Published online: 9 November 2004

Abstract In bone status assessment, proper quality assurance/quality control is crucial since changes due to disease or therapeutic treatment are very small, in the order of 2–5%. Unlike for dual X-ray absorptiometry, quality control procedures have not been extensively developed and validated for quantitative ultrasound technology, limiting its use in longitudinal monitoring. While the challenge of developing an ideal anthropometric phantom is still open, some manufacturers use the concept of the internal digital phantom mimicking human characteristics to check the stability of their device. The objective of the study was to develop a sensitive model of quality control suitable for the correction of QUS patient data. In order to achieve this goal, we simulated a longitudinal device lifetime with both correct and malfunctioning behaviors. Then, we verified the efficiency of digital phantoms in detecting those changes and subsequently established the in vitro/in vivo relationship. This is the first time that an attempt to validate an internal digital phantom has made, and that this type of validation approach is used. The digital phantom (DP) was designed to mimic normal bone (BUAP2) and osteoporotic bone (BUAP1) properties. The DP was studied using the UBIS 5000 ultrasound device (DMS, France). Diverse malfunctions of the UBIS-5000 were simulated. Several series of measurements were performed on both BUAP1 and 2 and on 12 volunteers at each grade of malfunction. The effect of each simulated malfunction on in vivo and in vitro results was presented graphically by plotting the average BUA values against the percentage change from baseline. The change from baseline in BUA was modeled using linear regression, and the in vivo/in vitro ratio was obtained from the model. All experimentations influenced the measure of BUAP1 and 2 as well as the measure of our 12 volunteers. However, the degree of significance varied as a function of the severity of the malfunction, and the results also differed substantially in magnitude between in vivo and in vitro. Indeed, the DP was about 10 times more sensitive to variations of the transfer function than was the in vivo measurement, which is very reassuring. The sensitivity of the digital phantoms was reliable in the determination of simulated malfunctions of the UBIS-5000. The digital phantoms provided an accurate evaluation of the acoustic performance of the scanner, including the fidelity of transducers. In light of these results, the QC approach of the UBIS-5000 will be extremely simple to implement compared with other devices. Indeed, since the digital phantom was automatically measured during every patient measurement, the QC approach could be built on an individual level basis rather than on an average basis.

Keywords Internal digital phantoms · Monitoring · Quantitative ultrasound device · UBIS 5000

D. Hans (✉) · I. Alekxandrova · G. Conicella · D.O. Slosman
Division of Nuclear Medicine, Geneva University Hospital,
1211 Geneva 14, Switzerland
E-mail: didier.hans@hcuge.ch
Tel.: +41-22-3727252
Fax: +41-22-3727255

C. Njeh
OARG, University of California San Francisco,
San Francisco, Calif., USA

C. Le-Floch
IRIS Servier, Courbevoie, France

P. Laugier
Laboratoire d'Imagerie Paramétrique UMR,
CNRS 7623, Paris, France

Introduction

Quality assurance (QA) in bone status assessment involves performance evaluation of the equipment and the operator in order to improve the reliability of the results. Quality control (QC), on the other hand, is an composite of sampling and statistical testing procedures designed to ensure adequate quality of the finished product. In bone status assessment, proper QA is crucial, since changes due to disease or therapeutic treatment are very

small, of the order of 2–5%. Sources of error include both measurement errors and equipment malfunction [1,2,3]. Van Daele et al. reported an upward trend over 3 years for SOS and a downward trend for BUA of an Achilles device, explained by a slight narrowing of the distance between the transducers over time [4]. Similarly, Krieg et al. mentioned in the Semof multicentre study that both the Achilles and the Sahara ultrasound devices shown significant fluctuations over time which probably affect the in vivo results. He reported that over a 2-year period, 21 interventions were performed on the ten Sahara devices (seven calibrations and 14 pads replacements), while on the ten Achilles devices there was still a significant trend of changes over time (adjusted for temperature) ranging from −0.8 to +0.6 dB/MHz per year for BUA, and from +2.1 to +5.1 m/s year for SOS [5]. Hans et al., in the Epidos multicentre study, showed significant differences between QUS devices of the same brand for the SOS up to 6 m/s in vitro and 8 m/s in vivo, but not for BUA, explained by the narrow dynamic range and the temperature dependence of the SOS [6]. Similar difference (up to 11 m/s for the SOS) has been also reported by Krieg et al. [5]. These magnitudes of difference correspond for some studies at the average difference between osteoporotic patients and age matched controls [7].

Equipment stability or malfunction can be monitored using a bone-equivalent phantom. Phantoms are important components of a QA program. In order to emulate the in vivo measurement as closely as possible in terms of geometric and acoustic properties, the phantom should ideally be anthropomorphic [2,8]. It is hoped that changes in the scanner which affect the in vivo measurements will also affect the phantom measurements in the same manner and thus be detected by the QC program, and that the variations of the in vivo results may be predicted from the phantom-based device calibration. As a result, one can correct patient measurements for the effect of an unstable device based on the drift or shift of the phantom measurements by applying a correction factor to the patient data. This procedure would guarantee that the patient data reflect the ''biological or therapeutic'' reality and are not affected by device malfunctions.

Appropriate QA and QC procedures have been well developed and validated for dual X-ray absorptiometry (DXA). These are now widely applied to multicenter clinical trials in order to monitor DXA devices used to assess treatment effects on bone mineral density [9,10]. However, this is not the case for quantitative ultrasound technology, for which almost no QC approach has been tested in clinical trials [11].

Anthropomorphic phantoms used for DXA are relatively simple in construction and the interaction between X-rays and bone-equivalent phantoms is well known. This is again not the case for quantitative ultrasound, for which the theory of propagation through cancellous bone is quite complex and not well understood. As a result, the manufacturing of an anthropomorphic phantom specific for ultrasound is very difficult and expensive. Presently, there are no universally accepted phantoms, but only ''manufacturer specific'' phantoms that are not anthropomorphic and whose daily changes may not reflect what occurs in vivo. In general, these phantoms consist of a piece of geometrically shaped, homogenous material designed to simulate the typical values of frequency-dependent attenuation and speed of sound encountered in vivo in cancellous bone, regardless of the actual geometry and spatial non-homogeneity of the skeletal tissue. These materials are well suited to help monitoring equipment drift or to detect electro-acoustic or electronic malfunctions (i.e. to detect any variations over time of the measuring device transfer function), although the reliability of these phantoms may be affected by external factors such as storage, aging and temperature [2,8]. The positioning of the phantom during the scan acquisition could also lead to monitoring errors due to possible non-homogeneity in the composition of the phantom material. Several non-ultrasound device manufacturers have also developed such external ''ultrasound specific'' phantoms which are still under evaluation for quantitative ultrasound (QUS) assessment, such as the Leeds QUS phantom and the Vancouver phantom [2,8,12]. The Leeds phantoms come in three varieties that have different QUS values in the clinical range. Although potentially interesting, these phantoms are very much temperature and positioning dependent. The Vancouver phantom, on the other hand, consists of a reticulated vitreous carbon form filled with heavy grade USP oil. Similar to the Leeds phantom, the Vancouver phantom comes in several varieties that have different properties. This phantom set has BUA in the clinically relevant range (24–70 dB/MHz), but velocity (1650 m/s) does not change with porosity as is seen in vivo. Positioning is less of an issue, but temperature adjustment is also needed.

To overcome the difficulties encountered in using external phantoms (instability, aging, poor spatial homogeneity), Langton has suggested using an internal electronic phantom consisting of an electronic filter attached to the receiving probe, and specifically designed to reproduce the frequency filtering effect of the cancellous bone on the received ultrasonic signal [13]. The concept of an internal phantom has been extended by Laugier, who proposed the use of internal digital phantoms [14]. Designed to test scanner performance for broadband ultrasound attenuation (BUA) measurement, Laugier's approach does not require any additional hardware, but simply uses the signal transmitted through a temperature-controlled water bath and combines several device responses obtained at various input voltages. According to Laugier, such phantoms are very stable as they do not require any manipulation during the measurement and are not influenced by external factors such as aging or temperature. The digital BUA phantom concept has already been implemented into one QUS device (UBIS 5000; DMS, Montpellier, France). However, the use of digital BUA phantoms to simulate both normal and osteoporotic bone properties

raises a series of questions: (1) are digital phantoms efficient in detecting any drift or shift in the machine? Consequently, could they be used to monitor the stability of QUS devices? (2) Since an internal digital phantom does not require a physical object between the transducers and by nature is not an anthropomorphic phantom, is the in vivo situation well simulated? (3) Is the hypothesis "the ratio variation in vivo/in vitro equals one" valid? (4) Could digital phantoms be used for longitudinal QC in clinical settings?

In order to validate a new phantom, one should ideally follow a QUS device over a long period of time with daily phantom and patient measurements, hoping that several malfunctions would occur. However, the feasibility of such a study is almost impossible since the patient's values change biologically over time and the well-characterized malfunctions are not very likely to occur frequently enough.

The objective of the study was to develop a sensitive model of quality control suitable for the correction of QUS patient data. In order to achieve this goal, we simulated a longitudinal device lifetime with both correct and malfunctioning behaviors. We then verified the efficiency of digital phantoms in detecting those changes and subsequently established the in vitro/in vivo relationship. This is the first time that an attempt to validate an internal digital phantom is made and that this type of validation approach is used.

Material and methods

We evaluated the behavior of digital phantoms in several situations by simulating different malfunctions of the UBIS-5000 ultrasound device.

UBIS 5000 description

The UBIS-5000 device is a water-based QUS system. It generates ultrasound images of the calcaneous using a pair of broadband-focused transducers (center frequency 0.5 MHz, diameter 29 mm, focus 50 mm) submerged in a water bath [14,15,16]. The device is equipped with a thermostat that maintains the water temperature at 30 ± 2°C. The measured parameter is the BUA. The ultrasonic attenuation in bone varies quasi-linearly with frequency in the commonly used frequency bandwidth (200–800 kHz). The BUA represents the slope of a linear regression fit to the measured frequency-dependent attenuation. The frequency-dependent attenuation is obtained from the ratio of the Fourier spectrum between a pulse transmitted through the bone and the Fourier spectrum of a reference pulse transmitted through water.

Digital phantom description

The digital phantom was designed to mimic normal bone (BUAP2) and osteoporotic bone (BUAP1) properties. The BUA values are registered for both digital phantoms during each in vivo or in vitro examination. The results are then displayed in the QC section of the UBIS software for each examination.

By applying different voltages (i.e. 200 V, 63 V, 20 V and 2 V) to the transmitting probe to measure the reference pulse through water, different attenuations can be simulated. Using 200 V as a reference, these voltages correspond to attenuations of 0, 10, 20 and 40 dB, respectively. As shown in Fig. 1, the Fourier spectra of signals transmitted through water without any external phantoms and using the voltages mentioned above have similar shapes but different amplitudes. As a consequence, the spectral differences obtained using the 200 V spectrum as a reference result in attenuation curves which are frequency-independent, as shown in Fig. 2. The amplitude spectrum of the different reference signals may be combined in order to simulate the attenuation effect as a function of frequency. For example, by combining the value of the attenuation obtained at 250 kHz with the voltage at 63 V and the attenuation at

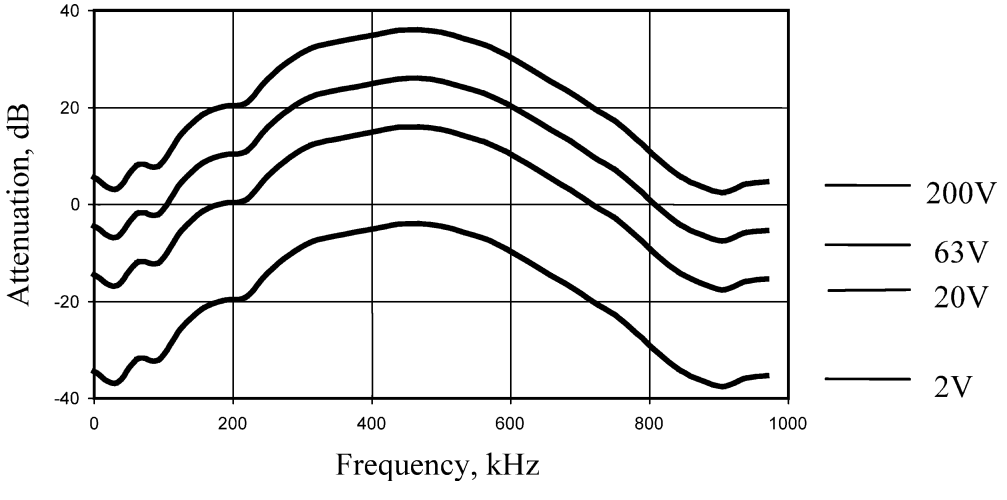

Fig. 1 Amplification curves: three attenuations can be calculated by using the voltages 63 V, 20 V and 2 V and considering the value of 200 V as a reference

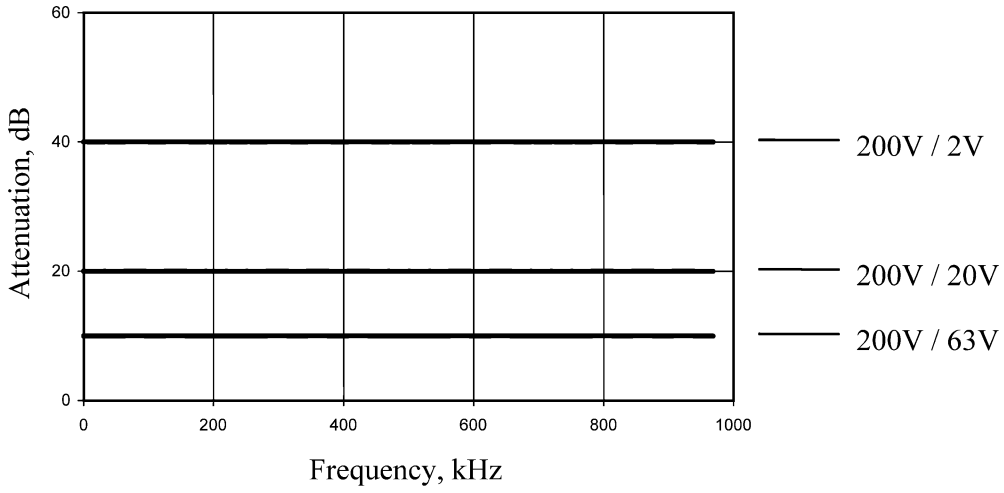

Fig. 2 Attenuation curves via the signal of 200 V

500 kHz with the 20 V voltage, and by calculating the slope of attenuation between those two frequencies, a slope of (20 dB−10 dB)/250 kHz = 40 dB/MHz is obtained. Figure 3 illustrates an example of this type of combination. The slope of the dotted curve is 40 dB/MHz. This value corresponds to the BUA of an osteoporotic woman. Similarly, combining the value of the attenuation obtained at 250 kHz with the voltage at 20 V and the attenuation at 500 kHz with the 2 V will result in a slope of 80 dB/MHz, which is representative of the average BUA for a young non-osteoporotic subject. As demonstrated, two different values of BUA were obtained from reference measurements in water with different input voltages. These two values are believed to represent normal and osteoporotic digital phantoms.

Studied population

Eight women and four men, aged from 28 to 72 years (mean age 41 ± 14.3 years), recruited within the employees or their relatives of the Nuclear Medicine Division of the Geneva University Hospital, were measured (two measurements for each time point) on the UBIS-5000 device in normal conditions and during the simulated malfunctions described hereafter. The inclusion of both sexes in our study ensured that the measured values covered a large range of BUA. None of the subjects had ankle edema or conditions or treatment known to affect bone metabolism. The characteristics of the population are presented in Table 1. Baseline measurements were taken between each set of malfunctions.

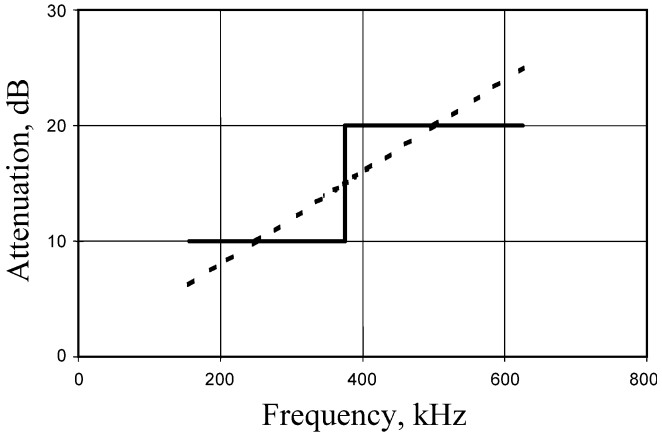

Fig. 3 Function of simulated attenuation: to obtain the BUA of 80 dB/MHz, corresponding to the average BUA value for a young non-osteoporotic population, the combination of 20 dB and 40 dB signals should be used

Simulated malfunctions

The malfunction program was established in collaboration with the maintenance department of the DMS company in order to mimic the most frequent breakdowns of the ultrasonic probes or acquisition electronics. The following situations were investigated: variation of the transfer function at reception and emission, modification of footbath temperature and malfunctioning transducers. The impact of simulated malfunctions was assessed in all subjects.

Variation of the transfer function at reception

The malfunction most frequently observed on the UBIS-5000 device was a deterioration of the ultrasound signal (e.g. caused by a broken probe). We therefore simulated a variation of transfer function at the reception stage. Figure 4 represents the system before modification, i.e. the standard commercial version. The reception stage is composed of two amplifiers, A1 and A2. The first amplifier A1 has a fixed gain of 6 dB. The second amplifier A2 has a variable gain between 0 and 60 dB. The high frequencies are then eliminated using a low-pass filter with a cutoff frequency of 4 MHz. The signal is then converted by an analog-to-digital converter (ADC), and transmitted to a PC-based computer for

Table 1 General statistics for baseline measurements

	In vitro			In vivo ($n = 12$)			Temperature
	BUAP1 (dB/MHz)	BUAP2 (dB/MHz)	BUA (dB/MHz)	Age (years)	Weight (kg)	Height(cm)	T (°C)
Mean value	39.13	81.50	61.56	41.0	57.75	166.17	31.5
SD	0.11	0.66	4.24	14.3	8.62	11.54	0.47
Min	38.83	79.71	53.00	28.2	47.00	151.00	30.4
Max	39.35	82.93	72.20	71.5	80.00	189.00	32.6

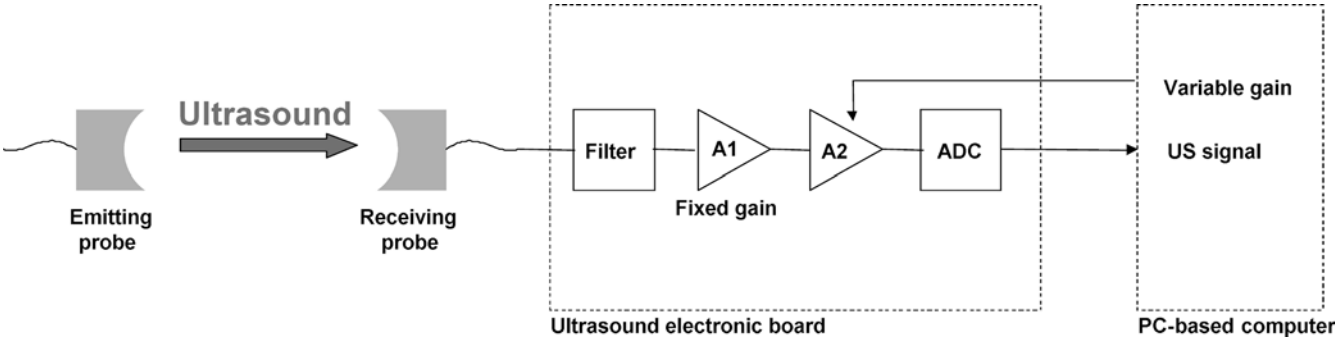

Fig. 4 Block diagram of system electronics: reception stage

digital signal processing. The computer then adjusts the gain of A2 in real time, so that the received signal amplitude always fits the dynamics of the ADC. Changing the gain of A1 (~6 dB) was made possible by replacing a fixed resistance of the ultrasound electronic board with a variable, manually adjusted resistance in order to achieve a gain from 6 dB to −80 dB. The signal amplitude level was controlled using the standard maintenance software provided with ultrasound device. This electronic modification of the system enables the simulation of a loss of sensitivity of the receiving stage (i.e. the receiving probe, A1, A2 and the filter): 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 56 dB.

Variation of transfer function at emission

The simulation of the transfer function variation at the emission stage was possible by addition of a variable resistance from 0 to 10 kΩ (Fig. 5). When the resistance is fixed at 10 kΩ, normal working conditions of the UBIS-5000 are obtained. The further the value of this resistance is decreased, the more the emitting voltage decreases, thus resulting in a reduction of the amplitude of the emitting ultrasound signal. Due to the sensitivity of the variable resistance, it was not possible to start with a mild amplitude variation. The signal amplitude loss ranged from −25 to −56 dB compared with the baseline, which was fixed at 0 dB.

Modification of footbath temperature

A switch on the UBIS hardware allowed a technician to modify the water temperature control from 27 to 33°C in steps of 2°C. We were able to increase the tested temperature range by adding hot or cold water (with surfactant). The obtained temperature range varied from 25 to 36.5°C.

Malfunctioning transducer

The pair of normally working transducers on the UBIS-5000 was successively replaced by two pairs of malfunctioning transducers, with a gain loss of about 8.5 dB and −36 dB, respectively. Apart from the probe, all the electronics remained unchanged. The results were compared with baseline data on variation of the transfer function at the reception stage. The real malfunction (malfunctioning transducers) was evaluated against the simulated malfunction (a variation of transfer function at the reception stage) in order to verify the validity of the malfunctioning probe model.

Statistical methods

The normality assumption of the dataset was evaluated using the JMP statistical package (version 3.2.2; SAS Institute Inc.). In vivo and in vitro data were then summarized using descriptive statistics (mean, standard deviation, minimum and maximum). These calculations were performed for baseline measurements and for each level of all simulated malfunctions. The effect of each simulated malfunction on in vivo and in vitro results was presented graphically by plotting the average BUA values against the percentage change from baseline. The change from baseline in BUA was modeled using linear regression, and the in vivo/in vitro ratio was obtained from the model.

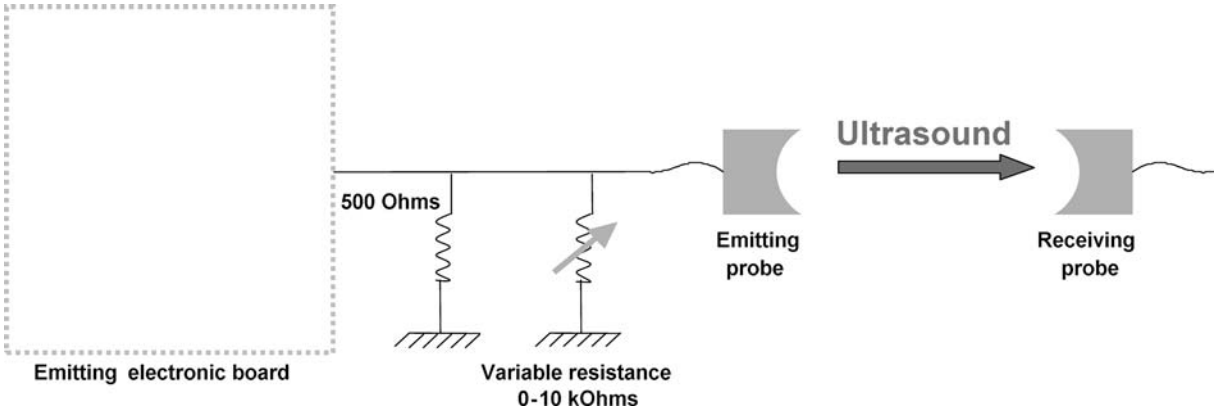

Fig. 5 Block diagram of the system electronics: emission stage

Results

Baseline measurements were collected for each volunteer, leading to a total of 76 baseline measurements for both digital phantoms, BUAP1 and BUAP2. The impact of the simulations of the ultrasound measurements is reported below.

Variation of transfer function at reception

At every modification of gain, each subject was measured twice and the mean BUA value was calculated. For each subject, the difference from baseline was then calculated and expressed in percentage. The results of the experiments are summarized in Table 2 and Fig. 6. The 95% confidence interval (CI) of the in vitro (1.7%) and in vivo precision (1.9%) is represented in Fig. 6. All variations of data outside of given limits are considered to be significant. BUAP2 reached a peak around 15 dB of amplitude loss, while BUAP1 reached its peak around 35 dB. The BUA (expressed in percentage) values then decreased, undershot and returned to the baseline level for BUAP1, while they remained lower than the baseline for BUAP2. A similar pattern was observed in vivo, where the BUA value increased with the decrease in gain and reached a peak at around 45 dB. The BUA value then decreased below the baseline level. When considering the in vitro and in vivo precision at 95% CI, the in vitro values reached a significant level quite rapidly (at 5 dB of gain loss), whereas none of the in vivo values reached the significance level before 35 dB of gain loss.

Variation of the transfer function at emission

We simulated malfunction of the emission electronics by altering the voltage of the signal driving the ultrasound emitter using a variable resistance. Due to the sensitivity of the potentiometer, we were not able to measure a loss in emission inferior to 25 dB with sufficient precision. The simulated malfunction, when applied in vivo, showed that a decrease in the voltage resulted in an increase in the BUA values (Table 3 and Fig. 7).

It appears that the digital normal phantom was less sensitive to a mild loss of signal amplitude than the osteoporotic one. The osteoporotic phantom increases

Table 2 Variation of transfer function at reception

Subtracted gain, dB	n[a]	Mean BUAdB/MHz			Mean SD				Difference to the baseline, %	
		P1	P2	In vivo	P1	P2	In vivo	P1	P2	In vivo
0	76	39.13	81.50	61.68	0.11	0.66	0.37	0.00	0.00	0
5	24	40.18	82.75	61.53	0.08	0.58	0.34	2.68	1.53	−0.25
10	24	40.63	86.18	61.87	0.05	0.71	0.30	3.85	5.75	0.31
15	24	41.28	89.10	61.95	0.12	0.57	0.31	5.50	9.33	0.44
20	23	42.59	87.12	62.13	0.17	0.50	0.18	8.84	6.90	0.73
25	24	43.44	84.33	62.17	0.10	0.42	0.33	11.02	3.48	0.78
30	24	45.67	80.55	62.38	0.31	0.52	0.21	16.72	−1.17	1.13
35	24	47.37	76.95	62.86	0.22	0.60	0.31	21.07	−5.59	1.89
40	24	43.89	76.85	63.25	0.40	0.85	0.48	12.16	−5.70	2.47
45	24	37.91	76.82	63.50	0.49	1.02	0.27	−3.11	−5.74	2.89
50	24	39.20	74.07	62.03	0.23	2.09	0.31	0.19	−9.11	0.52
56	21	39.23	74.49	59.82	0.33	2.42	0.36	0.27	−8.60	−3.06

[a]Total number of measurements per experimentation

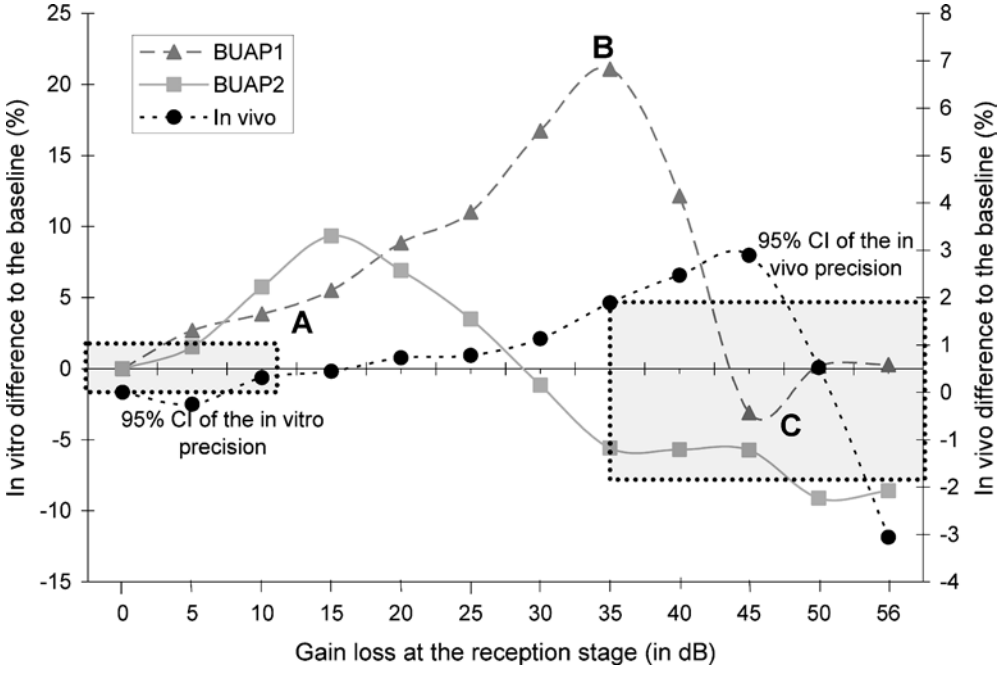

Fig. 6 Result of the gain loss at the reception stage on the in vivo and in vitro (digital phantoms) BUA measurements. Results are expressed as a percentage of baselines. The 95% confidence interval of the in vitro and in vivo precision is represented by dotted area

Table 3 Variation of transfer function at emission

Subtracted gain, dB	n[a]	Mean BUAdB/MHz			Mean SD			Difference to the baseline, %		
		P1	P2	In vivo	P1	P2	In vivo	P1	P2	In vivo
0	76	39.13	81.50	61.45	0.11	0.66	0.37	0	0	0
25	22	59.44	77.16	61.97	10.98	22.49	0.51	51.9	−5.3	0.5
30	24	60.78	67.00	62.26	10.79	23.89	0.44	55.3	−17.8	1.3
35	24	57.28	59.04	62.72	8.03	19.49	0.28	46.4	−27.6	2.1
40	24	50.90	48.29	62.16	4.81	18.45	0.24	30.1	−40.7	1.3
45	24	43.73	28.98	60.40	6.09	13.42	0.37	11.7	−64.4	−1.4
50	24	40.71	19.75	56.40	8.22	9.00	0.57	4.0	−75.8	−7.8
56	24	34.12	15.94	50.88	6.36	11.04	0.99	−12.8	−80.4	−16.7

[a]Total number of measurements per experimentation

its value to reach a peak around 30 dB and then decreases. In contrast, the normal phantom was stable until about 25 dB and then decreased with a pattern similar to the osteoporotic phantom. The beginning of the curve for the in vivo BUA is relatively flat, followed by a slight increase. This increase started at around 25–30 dB loss at the emitting stage. It is important to note that above 30–35 dB of loss, the UBIS software will trigger an alarm due to electronic fault. As a consequence, the effect of such a malfunction will be minimal in the authorized range of functioning.

In vivo/in vitro ratio

In order to analyse the loss of signal amplitude malfunction we used ultrasound values up to the amplitude peak of the osteoporotic digital phantom (−35 dB), because they follow a predictable pattern. The data beyond this point are affected by the limitations of the internal electronics of the UBIS-5000 and become useless for the analysis. Furthermore, an alarm will be triggered when a distortion of more than 30–35 dB is recorded by the system. As a consequence, we detected trends before the device reached a level of malfunction corresponding to this peak. Allowing trends to be detected after this point would imply that the measured values become unrecoverable.

The measurements on digital phantoms were found to be as sensitive to the simulated malfunction as the in vivo measurements. According to the correlation coefficient, 92% of the variation in vivo is explained by the in vitro simulation (Fig. 8). The in vitro/in vivo ratio was estimated to be about 1/10 for the BUA variation in amplitude at the reception stage, for both phantoms, and about 1/16 for the emission up to the peak of the BUAP1 phantom, corresponding about −35 dB. Nevertheless, it is difficult to have confidence in the regression slope (linking the in vitro and in vivo data in the simulation at emission) and to make a sensible conclusion based on insufficient number of time points.

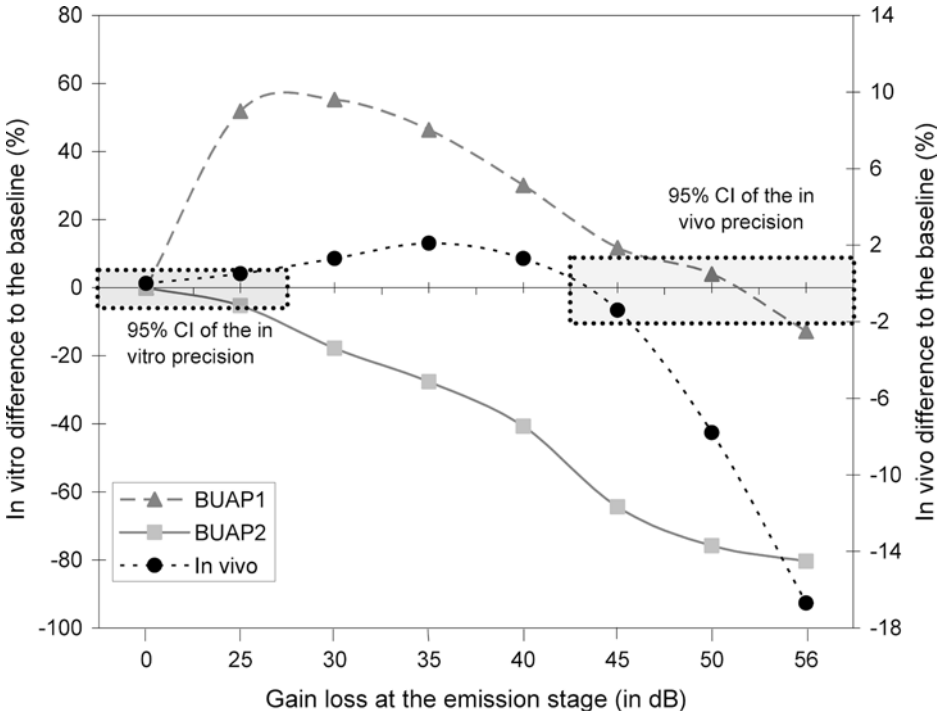

Fig. 7 Result of the gain loss at the emission stage on the in vivo and in vitro (digital phantoms) BUA measurements. Results are expressed as a percentage of baselines. The 95% confidence interval of the in vitro and in vivo precision is represented by the *dotted* area

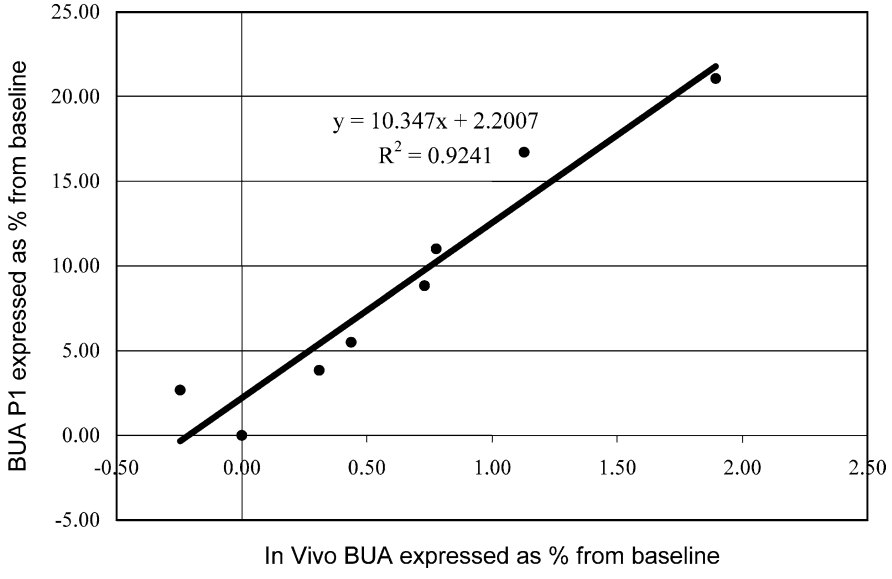

Fig. 8 The ratio BUAP1 value versus in vivo average value calculated from the experience of gain at the reception stage. The ratio enables us to estimate the sensitivity of the digital phantom in detecting malfunction compared to real subject data and to apply the appropriate correction factor to in vivo data based on the in vitro results

Modification of footbath temperature

According to the regression slope (regression not shown), a temperature variation of 1°C corresponds to an increase in the BUAP1 value by 0.012 dB/MHz, that is 0.03% (and 0.045 dB/MHz ~0.05% for BUAP2). Since the device is temperature controlled at 30°C ± 2°C (out of this range, an alarm will be triggered), the total BUAP1 variation for the temperature change of 4°C will be equal to 0.046 dB/MHz or 0.11% (and 0.18 dB/MHz or 0.20% for the BUAP2). That value is 3–6 times less than the in vitro precision for the digital phantom.

Malfunctioning transducer

We verified the validity of the malfunction's simulation in the case of a loss of signal amplitude at the reception stage, in vivo as well as in vitro, by creating a real breakdown using transducers with a known loss of sensitivity. The two broken transducers showed a loss of sensitivity of −10 dB and 36 dB, respectively. The results of the real breakdown were expressed as a percentage with respect to baseline, and were then compared with theoretical simulated values (Table 4). The simulated malfunction and the real breakdown produce the same tendency and a similar magnitude in BUA changes.

Table 4 In vivo versus in vitro comparison: malfunctioning transducer in percentage of baseline value. Values are given as percentages

	BUA P1	BUA P2	In vivo BUA
Gain loss of 36 dB			
Simulated malfunction	16.5	−4.0	2.9
Real breakdown	17.5	−9.5	3.1
Gain loss of −10 dB			
Simulated malfunction	3.8	4.5	0.3
Real breakdown	4.0	3.0	0.3

Discussion

The internal digital phantom allows automatic calibration of the device. The advantage of the digital phantom is its ease of use, since it does not require any manipulations during the measurement, thus avoiding positioning errors. In addition, the influence of external factors on digital phantoms, such as aging and temperature, should be less pronounced than on external phantoms. As a consequence, internal phantoms are expected to have a much smaller drift than external phantoms in short- and long-term use.

We studied the impact of malfunction simulations on BUA values by simultaneously measuring the calcaneous of 12 volunteers and the digital phantom. The in vivo group was first divided into two subgroups aiming to represent osteoporotic and normal populations (subjects with low values and normal values), but since no differences were found either in the ultrasound signal pattern or in the results, we decided to pool all subjects' values in the same group. Overall, our data are very comparable with results from a previous exploratory study [17,18], suggesting the reproducibility of our experiment.

The impact of the transfer function variation in the receiving signal exists on the in-vivo ultrasound parameters, and is overall directly proportional to the malfunction. However, this impact barely reaches a significant level compared with the in-vivo precision of the QUS, and after a substantial deterioration of the signal (at least −25 dB). Moreover, we can clearly see a deterioration of the image quality around 40 dB of amplitude loss or greater. In addition to the alarm set at −35 dB, the visual quality control of the examination should enable the capture of such images and classify them as poor quality. It is very interesting to observe that in vivo and in vitro results show an identical curve shape when simulating the malfunction of loss of signal amplitude. We observed a parallel increase in BUA and gain until the maximum peak and a steep decrease after that point. The attenuation compares the signal amplitude that passes through the reference material (water) to the signal amplitude that passes through the sample (bone), and is expressed in dB. The BUA is the slope of the linear regression over the attenuation curve as a function of frequency. The same principle is applied to digital phantoms, but the bone behavior is simulated by the device itself. During the simulated malfunction, we reduced the transfer function by modifying the fixed gain of the amplifier A1 (Fig. 4), thus changing the attenuation. In the example of the BUA curve of the osteoporotic phantom (Fig. 6), several regions corresponding to different parts of this simulation, can be observed: (1) the first part of the curve (up to point A) shows the behavior of the device with an unmodified or slightly modified gain; (2) the section between points A and B in the curve illustrates how the UBIS-5000 reads a progressively increasing BUA as the gain of the amplifier is reduced (impacting the transfer function); (3) when the gain in amplifier A1 is reduced, we also reduce the amplitude of the signals (in water and through bone) at the output. At point B, the signal becomes extremely weak and at high frequencies this signal is even weaker than the background noise. The more the gain of A1 is reduced, the larger will be the portion of the signal curve that is covered by the background noise. Since the noise amplitude is larger than the signal amplitude in this area, the measurements will give the false impression that the signal is less attenuated than it really is. Consequently, the calculated attenuation curve will become progressively flatter, resulting in a decreased BUA, which appears between B and C; and (4) at point C, the gain of A1 is so reduced that A2 is set to its maximum gain by the software. A further decrease in the gain of A1 cannot be compensated by an increased gain in A2, resulting in a decreased signal after the amplification chain (A1 + A2). The signal measured at this point is probably perceived as a weaker signal through bone, leading to the interpretation that the bone shows a higher attenuation. Consequently, the measured BUA increases with a decreased gain in A1 (behavior after point C).

The impact of loss of amplitude of the emitting signal is not as clear as the previous experiment (reception stage). Nevertheless, there is an influence of the emitting signal on the in vivo ultrasound parameters, which also does not reach significance compared to the in vivo precision.

Based on our experiment, it appears that the temperature fluctuation of the water bath in the working range of the UBIS-5000 does not influence ultrasound measurements in vitro. This result is not surprising, since at the level of frequencies used to generate these ultrasounds, the attenuations in water are very small. In comparison with the most commonly used external phantoms, Njeh has demonstrated that an increase in temperature of 1°C caused an increase of 0.1-0.4% in BUA in the Leeds phantom, and an increase of 0.1-1.6% in the Vancouver phantom [19]. Krieg has observed with the Achilles black rubber phantom that for each increase of 1°C, BUA decreased by −0.6 dB/MHz, and SOS by −4.3 m/s ($P < 0.001$ for all three correlations). In our study, a temperature variation of 1°C causes the BUA to increase by 0.03% with the osteoporotic phantom and by 0.05% with the normal phantom.

The in vitro/in vivo ratio of 1–10.35 (or even more for the speed of sound ~60.54) plays a favorable role

in the QC approach and highlights the higher sensitivity of the phantom in detecting breakdowns. The digital phantoms identify even small variations of gain prior to the in vivo observation. That means that any shift or drift of the apparatus, at least during loss of signal amplitude on the reception (or emission at a lesser extent) stage malfunction, will be registered by digital phantoms before the failure of the machine influences the in vivo results. Indeed, when considering an in vivo precision of 0.66% for BUA, to have confidence in the impact of the malfunction in vivo at 95% CI (i.e. about 1.9% for the BUA), one must observe about 20% ($\sim$10.35×1.9) deviation from baseline in vitro for BUA. The early detection of device malfunction will allow the system to be repaired immediately, or for a retrospective correction of the in vivo data to be made by applying the correction factor. Therefore, the patient results will be due to biological or therapeutic effect and not due to incorrect device performance. This is the first time that the ratio between in vivo and in vitro in QC application has been investigated.

The favorable results of the digital phantoms have two possible explanations: (1) the good linearity of the electronics in the working range, and (2) the digital phantoms are working at a lower voltage (63 V, 20 V and 2 V) than in vivo measurements (200 V). Small variations at the emission or reception will be more easily captured if the emission voltage is low (2 V). This explains why the osteoporotic digital phantom is more sensitive to malfunction.

Overall, when considering the loss of amplitude in emission or reception of up to 35 dB (beyond that point an alarm is triggered), none of the experiments significantly influenced the in vivo results (when considering the 95% CI of the in vivo precision). As a result, the stability of the device is well captured by the digital phantom with minimal in vivo impact within the monitored range. Out of the two BUA parameters generated by the digital phantom, BUAP1 corresponds better to the usual studied populations in clinical trials and is more sensitive to emission and reception malfunctions. Nevertheless, the BUAP2, the pattern of which differs between the two experiments, could be used to differentiate the type of malfunction and the subsequent type of correction.

There are some limitations to our study. Indeed, this work is relevant only when applied to the use of the UBIS-5000 devices, since some aspects of the study depend on limitations derived from the device's electronics design. These limitations, depending on engineering choices, can be different from one type of device to another. Hence, some reasoning valid for the UBIS-5000 might be erroneous when applied to a device from a different manufacturer. Although SOS is measured by the UBIS devices, it is not reported by the manufacturer on the patient printout. Furthermore, the digital phantom does not measure the SOS parameter. As such, it was decided not to study the SOS in this experimentation. Nevertheless, the water SOS could be also be used to monitor the stability of the device if needed, as was suggested by Hans et al., who co-developed the methodology for the Achilles device [11]. A similar approach should be applied for each individual ultrasound device of different brands, because their own characteristics may differ from UBIS-5000. In addition, we studied only the most common malfunctions and did not cover the whole spectrum of all possible malfunctions. Even though there are limitations in our approach, the results indicate that the sensitivity of the digital phantoms is reliable in the monitoring of device instability and malfunctions. As a result, one could imagine that all water based devices could use this digital phantom. For dry systems, the digital phantom developed by Langton et al. [13] could be an alternative, although it would need to be verified and tested.

In conclusion, the sensitivity of digital phantoms is reliable in the determination of simulated malfunctions of the UBIS-5000. The digital phantoms provide an accurate evaluation of the acoustic performance of the scanner, including the fidelity of transducers. In light of these results, the QC approach of the UBIS-5000 will be extremely simple to implement compared with other devices. Indeed, since the digital phantom is automatically measured during each patient measurement, the QC approach could be built on an individual level basis rather than on an average basis.

Acknowledgements Special thanks are due to B. Fournier and M. Julliet for their scientific input. We are also grateful to all our colleagues who volunteered to participate in this study. This study was funded by the Institute of International Research Servier (IRIS), the company Diagnostic Medical System (DMS) and the Swiss National Fund (subside no. 31-61994.00).

References

1. Blunt B, Jones D, Svensen R, Hans D, Feinblatt JD, Genant H (1998) Good clinical practice and audits for dual X-ray absorptiometry and X-ray imaging laboratories and quality assurance centers involved in clinical drug trials, private practice and research. J Clin Densitom 1:323–337
2. Fuerst T, Njeh CF, Hans D (1999) Quality assurance and quality control in quantitative ultrasound. In: Njeh CF, Hans D, Fuerst T, Glüer CC, Genant HK (eds) Quantitative ultrasound: assessment of osteoporosis and bone status. Martin Dunitz, London, pp 163–175
3. Orwoll ES, Oviatt SK, Biddle JA (1993) Precision of dual-energy X-ray absorptiometry: development of quality control rules and their application in longitudinal studies. J Bone Miner Res 8:693–699
4. van Daele PL, Burger H, De Laet CE, Hofman A, Grobbee DE, Birkenhager JC, Pols HA (1997) Longitudinal changes in ultrasound parameters of the calcaneus. Osteoporos Int 7:207–212
5. Krieg MA, Cornuz J, Hartl F, Kraenzlin M, Tyndall A, Hauselmann HJ, Lippuner K, Rizzoli R, Buche D, Theiler R, Dambacher MA, Neff M, Pancaldi P, Tanzi F, Wimpfheimer K, Burckhardt P (2002) Quality controls for two heel bone ultrasounds used in the Swiss Evaluation of the Methods of Measurement of Osteoporotic Fracture Risk Study. J Clin Densitom 5:335–341

6. Hans D, Schott AM, Chapuy MC, Benamar M, Kotzki PD, Cormier C, Pouilles JM, Meunier PJ (1994) Ultrasound measurements on the os calcis in a prospective multicenter study. Calcif Tissue Int 55:94–99
7. Njeh CF, Hans D, Li J, Fan B, Fuerst T, He YQ, Tsuda-Futami E, Lu Y, Wu CY, Genant HK (2000) Comparison of six calcaneal quantitative ultrasound devices: precision and hip fracture discrimination. Osteoporos Int 11:1051–1062
8. Strelitzki R, Clarke AJ, Truscott JG, Evans JA (1996) Ultrasonic measurement: an evaluation of three heel bone scanners compared with a bench-top system. Osteoporos Int 6:471–479
9. Lu Y, Mathur AK, Blunt BA, Gluer CC, Will AS, Fuerst TP, Jergas MD, Andriano KN, Cummings SR, Genant HK (1996) Dual X-ray absorptiometry quality control: comparison of visual examination and process-control charts. J Bone Miner Res 11:626–637
10. Wahner HW, Looker A, Dunn WL, Walters LC, Hauser MF, Novak C (1994) Quality control of bone densitometry in a national health survey (NHANES III) using three mobile examination centers. J Bone Miner Res 9:951–960
11. Hans D, Wacker W, Genton L, Paris E, Le-Floch C, Slosman D (2002) Longitudinal quality control methodology for the quantitative ultrasound Achilles+ in clinical trial settings. Osteoporos Int 13:788–795
12. Njeh C, Chen M, Fan B et al. (2000) In vitro (phantom) cross calibration of quantitative ultrasound devices. Osteoporos Int 11:14
13. Langton CM (1997) Development of an electronic phantom for calibration, cross-correlation and quality assurance of broadband ultrasound attenuation measurement in the calcaneus (abstract). Osteoporos Int 7:309
14. Laugier P, Fournier B, Berger G (1996) Ultrasound parametric imaging of the calcaneus: in vivo results with a new device. Calcif Tissue Int 58:326–331
15. Njeh CF, Blake GM (1999) Calcaneal quantitative ultrasound: water-coupled. In: Njeh CF, Hans D, Fuerst T, Glueer C-C, Genant HK (eds) Quantitative ultrasound: assessment of osteoporosis and bone status. Martin Dunitz, London, pp 109–124
16. Laugier P, Giat P, Berger G (1994) Broadband ultrasonic attenuation imaging: a new imaging technique of the os calcis. Calcif Tissue Int 54(2):83–6
17. Alekxandrova I, Slosman D, Hans D 2000 Quality control in ultrasound: electronic phantom versus in vivo measurements. Osteoporos Int 11:14
18. Alekxandrova I (2000) Effectiveness of digital phantoms for monitoring the stability of a quantitative ultrasound device in clinical setting Nuclear Medicine Division. Medical Thesis no. 10174 submitted at Geneva University, Geneva, pp 122 (supervised by Dr. D. Hans)
19. Njeh CF, Chen M, Fan B, Fuerst T, Wu C, Diessel E, Hans D, Genant H (1999) The impact of temperature on quantitative ultrasound: an in-vitro study using phantoms. J Bone Miner Res 14:S500