Reliability of Measurement of Skin Ultrasonic Properties *in Vivo*: A Potential Technique for Assessing Irradiated Skin

Y.P. Huang [1], Y.P. Zheng [1], S.F. Leung [2], A.F.T. Mak [1]

[1] Department of Health Technology and Informatics, Hong Kong Polytechnic University, Hong Kong, China

[2] Department of Clinical Oncology, Chinese University of Hong Kong, Hong Kong, China

Corresponding author:

Dr. Yongping Zheng

Department of Health Technology and Informatics

Hong Kong Polytechnic University

Hong Kong, China

Phone: 852-27667664

Fax: 852-23624365

Email: ypzheng@ieee.org

Abstract

Background/aims: Quantitative and objective technique to assess radiation-induced tissue fibrosis is important for clinicians to estimate the efficiency of radiotherapeutic schemes. It has been widely reported that ultrasonic properties are sensitive to changes of acoustic scatterers in biological tissues. Therefore, measurement of ultrasonic properties may serve as a potential assessment technique for irradiated skins. The aim of the present study is to investigate the reliability of such measurement so as to evaluate its potentials for future clinical applications.

Methods: Ultrasonic parameters including attenuation slope (β), integrated attenuation (IA) and integrated backscatter (IBS) were measured for the frequency range of 10 to 25 MHz from echographic signals of the forearm and neck dermis of 20 normal subjects *in vivo*. The intra- and inter-rater reliability of measurement was assessed in 10 normal subjects using intra-class correlation coefficients (ICC) and Bland-Altman test.

Results: The intra- and inter-rater measurement was demonstrated to be reliable as indicated by high ICC values generally larger than 0.80. In addition, the ultrasonic parameters could successfully differentiate the skins in the neck and forearm regions.

Conclusion: The measurement provided reliable information on the ultrasonic properties of the skins and could be potentially applied to comparative clinical trials to assess the late effects of radiotherapy on skins.

Keywords: radiotherapy – skin – fibrosis – high frequency ultrasound – ultrasonic tissue characterization – ultrasonic properties – reliability.

Introduction

Fibrosis is a common late side effect of radiotherapy. Assessment and quantification of the severity of tissue fibrosis are important in determining the curing ratio of different radiotherapeutic regimes or assessing the efficiency of various anti-fibrotic medicines (1). Although some standard protocols such as the LENT/SOMA scoring system have been established to assess the degrees of late radiation toxicities (1), clinical practice is still mainly limited to hand palpation. Quantitative and objective methods have not been intensively investigated for this purpose.

Being noninvasive and inexpensive, high-frequency (larger than 10 MHz) ultrasound has been successfully introduced in dermatology to detect the delicate skin structures and lesions in the last two decades (2-4). Applications of ultrasound reported in dermatology generally include two kinds of measurement, i.e., skin thickness and echogenicity. Skin thickness can be directly measured from the ultrasound images, while a common way to calculate the echogenicity is based on counting the number of pixels with a certain range of intensity values in images (5). This method has been used for the study of skin reactions induced by therapeutic or accidental radiation and high frequency ultrasonic imaging is recommended as a reliable tool to assess the irradiated effects on the skin (6, 7).

On the other hand, it has been widely reported that ultrasonic propagation properties are sensitive to the structural, compositional and pathological changes of acoustic scatterers in biological tissues (8). The propagation parameters are extracted from the radiofrequency (RF) signals, thus capable of providing new information on tissue conditions, such as frequency-dependent attenuation, in addition to the envelope signals used in the conventional ultrasonic imaging. Previous studies had showed that it was potential to monitor the skin changes induced by healing process or patch-test (9, 10). We have successfully used an ultrasound indentation technique to measure the mechanical properties of neck tissues with radiotherapy-induced fibrosis (11-13). During the study, we noted that the ultrasound signals scattered from the skin and subcutaneous tissues were different for subjects with different degrees of tissue fibrosis. In this study, we aimed to test the reliability of the high frequency ultrasound measurement for the skin assessment so as to prepare for the future clinical applications of this technique.

There had been some reports on the reliability of measurement of skin ultrasonic parameters in literature. For example, Guittet et al. (14) reported a coefficient of variation of 12% for the attenuation slope of skin; Fournier et al. (15) reported a standardized coefficient of variation of less than 20% for the backscatter spectral parameters of skin. However, it is difficult to transfer the measurement precision from one system to another due to the use of different devices and detection techniques. Accordingly, we investigated the reliability of measurement of ultrasonic parameters including attenuation slope, integrated attenuation and integrated backscatter in the skin *in vivo* using a 20 MHz ultrasonic imaging system. Intra-class correlation coefficients (ICC) and Bland-Altman test were used to estimate the reliability with respect to the intra- and inter-rater operations.

Methods

Data acquisition system

The system used for data collection was a 20 MHz ultrasonic imaging system developed for skin applications (Ultrasons Technologies, Tours, France). A handheld cylindrical probe, inside which there was a focused mono-element transducer, was used to collect echographic RF signals. Using the reflection signals from a planar steel plate immersed in a water tank, the measured central frequency of the transducer was 15 MHz with a -3 dB bandwidth of 10~25 MHz. A rigid removable cover with a flat membrane of 15 μm in thickness was tightly attached to the end of the probe to allow the ultrasound propagating through and to form the chamber for the acoustic coupling medium. The ultrasonic beam focused at approximately 2 mm beyond the membrane. Water was used to fill the cavity formed by the cover and the transducer. The transducer was translated linearly in the probe during operation. The contact area between the cover and the skin was small with a dimension of $11 \times 19\ mm^2$ and the effective width of image was 6 mm in the lateral direction. Each collected image provided a set of 256 RF signal lines, with 1024 data points in each line sampled at 100 MHz. The lateral resolution of the transducer was approximately 0.2 mm, therefore producing totally 28 independent lines per image, which were used for spatial averaging. In the studied region of interest (from 0.3 to 1.3 mm beneath the skin surface), we found that the average Pearson correlation for two adjacent lines was 0.24 ± 0.04 in the forearms of 5 normal subjects, indicating good independence for the signals.

Extraction of ultrasonic parameters

In this study, the skin was assumed to be composed of epidermis and dermis with a total thickness larger than 1.3 mm and the speed of sound was assumed to be 1.58 mm/μs throughout the whole skin layer (16). The ultrasonic propagation parameters were extracted from the spatially averaged RF signals using a multi-narrowband algorithm (17). The reference signals from a planar steel plate were collected in order to correct the system-dependent effects. The detailed procedure was described as follow.

Attenuation slope (β) and integrated attenuation (IA). A Hamming-window with a length of 50 points (0.40 mm) was used to gate the original temporal signal from the point 40 (0.32 mm) after the skin entry echo to the point 165 (1.30 mm) with 50% overlapping between two consecutive windows. This region was defined as the region of interest (ROI) as shown in Fig 1.a and 1.b. The skin entry echo was detected automatically in a custom-designed program based on the sudden change of the signal energy. According to the window selection principle, totally six short signals were obtained from each scanned line. Fast Fourier transform (FFT) was used to obtain the power spectra of the short signals after they were zero-padded to 512 points. The power spectra from the 28 independent lines were then averaged before they were divided by those of the corresponding reference signals to correct the system-dependent effects. The spectra of reference signals were obtained by applying the same window to the signals completely reflected from the steel plate in water at the same distance as the selected short signals. The signal lines which had saturated in the ROIs were excluded in the averaging procedure in order to avoid the contamination of spectra induced by the saturation. After correction, the six power spectra were logarithmically transformed (Fig 1.c) and regressed by the propagation distances of ROIs. Based on a commonly used hypothesis that the skin backscatter was homogeneous in all selected skin depths (18), the attenuation coefficients $\alpha(f)$ (unit: dB/mm) in 10~25 MHz were obtained (Fig 1.d). It was assumed that $\alpha(f)$ was linearly dependent on frequency, i.e.:

$$\alpha(f) = \beta \cdot f + \alpha_0 \qquad (1)$$

where β (unit: dB/mm/MHz) is the attenuation slope and α_0 (unit: dB/mm) is the intercept at zero frequency. β was then calculated from the regression of $\alpha(f_i)$ to frequency f_i obtained from the experiment. The integrated attenuation (IA, unit: dB/mm) was defined as:

$$\mathrm{IA} \equiv \frac{1}{f_h - f_l} \int_{f_l}^{f_h} \alpha(f) df \tag{2}$$

where $f_l = 10\,\mathrm{MHz}$ and $f_h = 25\,\mathrm{MHz}$, are the lower and upper bounds of the -3 dB bandwidth of the transducer.

Integrated backscatter (IBS). The signal in the ROI from the point 40 to 165 (0.32 to 1.30 mm) after the skin entry was used for the calculation of the backscatter coefficient. The temporal signals were gated using a Hamming-window (125 points) and zero-padded to 512 points before FFT transformation to power spectra. The power spectra of the independent lines of each image were averaged and divided by those of the reference signals using the same Hamming-window. The system-corrected spectra were then logarithmically transformed to obtain the backscatter spectra $B(f)$. The integrated backscatter (IBS, unit: dB) was then defined similarly to that of IA:

$$\mathrm{IBS} \equiv \frac{1}{f_h - f_l} \int_{f_l}^{f_h} B(f) df \tag{3}$$

where f_h and f_l are the range of the -3 dB bandwidth as in Eq. (2). It should be noted that IBS was not corrected by attenuation of the skin-film-water interface and the partial skin on top of the selected region. This parameter represented an average level of backscatter of the selected skin.

Tests of reliability and experiments on normal subjects

The *in-vivo* study was conducted for 20 normal subjects (age: 27 ± 3 yrs) under room temperature at approximately 25°C. In order to test the intra-rater reliability of this measurement, tests on 10 normal subjects were conducted by the first operator twice with a time lapse of one week. To test the inter-rater reliability, the measurement was also conducted by a second operator during the first data collection by the first operator. The two operators were both working in the ultrasound field and knew well the principles and details of the measurement. The selected testing sites were located on the palmar side (3 cm away from the radiocarpal joint and along the media nerve) and the dorsal side (in the middle portion of the forearm and 2 cm proximal to the ulnar styloid) of the distal forearm and also on both sides (5 cm below the mastoid bone) of the neck. We were interested in neck sites because we would recruit patients who had skin fibrosis in the neck due to radiotherapeutic radiation in future studies. The forearm testing sites served as the comparative parts of the neck sites in the same subject. The study was approved by the Research Ethics Committee of the investigators' affiliated institutions and written informed consent was obtained from each subject at recruitment.

The subjects were asked to be seated when the forearm sites were tested, and to lie on their sides when the neck sites were tested, both in a relaxing state. The probe was applied to the skin surface using a very small pressure in order to avoid changing the ultrasonic properties inconsistently by the probe pressure (15). Water was used as coupling medium between the probe and skin. Three repeated tests were conducted at each testing site and the averaged results of the three tests were used to represent the ultrasonic properties at that site.

Data analysis methods

The intraclass correlation coefficient (ICC) and the Bland-Altman test (19) were used to estimate the reliability of the ultrasound measurement. ICC represents the proportional contribution of variance of interest with respect to total variance

(variance of interest + measurement error). In our study, intra-rater and intra-rater reliability was respectively assessed by ICC(1,3) and ICC(3,3), the detailed meaning of which was described in (19). Bland-Altman test is used to analyze the individual difference of two repeated measurements. The reliability coefficient, as defined to be two standard deviations of the individual difference by assuming the mean difference to be zero, stands for the range of individual difference observed for 95% pairs of measurements (19). ICC and Bland-Altman test were widely used to quantify the reliability of measurement for continuous data.

The ultrasonic parameters in normal subjects were compared using the paired t-test between two sites in the forearm, in the neck, and between the forearm and neck regions. The parameters of two testing sites in the forearm or neck were pooled to be mean values of the forearm or neck if no significant difference was found in the two regions. All the statistical analyses were conducted by using the statistical software SPSS (SPSS Inc., Chicago, IL, USA). A comparison with $P < 0.05$ was used to indicate a significant difference.

Results

Table 1 shows the results of ICC and their 95% confidence intervals for the intra- and inter-rater measurement. All ICCs were larger than 0.80, which indicated a high reliability. Table 2 shows the results of Bland-Altman test for the individual difference of intra- and inter-rater operations. 95% CIs of the individual difference of ultrasonic properties included zero value, indicating the intra- and inter-rater measurement was unbiased.

Fig 2 shows the typical images obtained from the forearm and neck regions of a normal subject. It was consistently observed from these images that the neck skin in the dermal region was darker in comparison with the forearm skin. Fig 3 shows a summary of the ultrasonic properties measured from the 20 normal subjects. For all the parameters, no significant difference was demonstrated between two sites of the forearm ($P > 0.05$) and between two sides of the neck ($P > 0.05$). Therefore, data of the two sites within each region were pooled for the subsequent data analyses. For the attenuation coefficients, a linear relationship versus frequency with a regressed correlation coefficient larger than 0.90 could be generally obtained (Fig 1.d), showing that a linear relationship was good enough to describe the frequency-dependent characteristics. In the neck and forearm regions of the normal subjects, β was 0.328 ± 0.037 and 0.375 ± 0.037 dB/mm/MHz ($P < 0.001$), IA was 2.17 ± 1.90 and 0.58 ± 1.34 dB/mm ($P = 0.001$), and IBS was –33.55 ± 3.12 and –31.42 ± 2.44 dB ($P < 0.001$), respectively. The finding of a smaller IBS in the neck region was consistent with the observation of the dermal echogenicity in the images.

Discussion

A hand-held 20 MHz ultrasound probe was introduced in the current study to measure the ultrasonic propagation properties of skin *in vivo*. The reliability of intra- and inter-rater measurement was investigated in order to assess its potential use in clinical trials. Tests for 20 normal subjects were conducted and results were reported.

The results showed that all ICCs for the intra- and inter-rater tests were larger than 0.80, indicating that the measurement was highly reliable (20). It means that the influence of day-to-day and rater-to-rater variations were limited in the given test conditions. However, it should be noted that in order to maintain such high test reliability, basic training on the test operations for a rater was necessary before the rater could conducted the corresponding research. The key factors in operating the probe included gentle contact pressure, accurate probe orientation and good coupling between the probe and skin. If these conditions could be guaranteed, the measurement could be highly reliable and this ultrasound method was potential to be applied in clinical trials to assess the skin effects induced by radiotherapy.

The acoustic properties of the skins in the forearm and neck were demonstrated to be significantly different by the ultrasonic measurement. The difference was regarded as a reflection of the anatomic distinction of the skins in these two regions. Raju et al. (18) reported a larger β in the dermis of fingertip in comparison with that of two locations in the forearm. They attributed the difference to the potential variations of skin tension and collagen structure between the fingertip and forearm skins. A significantly smaller value of IBS was found in the neck than forearm region in the current study. Relevant study about skin echogenicity by Pellacani and Seidenari (21) demonstrated that the facial skin was much less reflective than the forearm skin. Thus the current findings were not contradictive to their results if the neck skin was assumed to be more similar to the facial skin in terms of ultrasonic scattering. Further investigations are required to explain the exact reason for the differences between the two regions.

Conclusion

The reliability of the current measurement of the ultrasonic properties was high with ICCs larger than 0.80. The ultrasonic parameters could discriminate the skins in the forearm and neck regions in the normal subjects. Considering that the operation of the current method was reliable, easy, and non-invasive, we expected that it would be potentially useful for the comparative clinical trials to assess the late effects of radiotherapy to the skin and future research towards this direction is being conducted.

Acknowledgement

This work was supported by the Research Grant Council of Hong Kong (PolyU 5245/03E) and the Hong Kong Polytechnic University. Sincere thanks were also given to Mr. Alex Choi Pongchi, Mr. Huang Qinghua, Dr. Chenxin, and Miss. Lu Minghua for their help in conducting some parts of the experiments.

References

1. Davis AM, Dische S, Gerber L, Saunders M, Leung SF, O'Sullivan B. Measuring postirradiation subcutaneous soft-tissue fibrosis: state-of-the-art and future directions. Semin Radiat Oncol 2003; 13:203-13.
2. Altmeyer P, el-Gammal S, Hoffmann K, eds. Ultrasound in Dermatology. Berlin: Springer-verlag, 1992.
3. Jemec GB, Gniadecka M, Ulrich J. Ultrasound in dermatology. Part I. High frequency ultrasound. Eur J Dermatol 2000; 10:492-7.
4. Rallan D, Harland CC. Ultrasound in dermatology--basic principles and applications. Clin Exp Dermatol 2003; 28:632-8.
5. Seidenari S, Giusti G, Bertoni L, Magnoni C, Pellacani G. Thickness and echogenicity of the skin in children as assessed by 20-MHz ultrasound. Dermatology 2000; 201:218-22.
6. Gottlober P, Kerscher MJ, Korting HC, Peter RU. Sonographic determination of cutaneous and subcutaneous fibrosis after accidental exposure to ionising radiation in the course of the Chernobyl nuclear power plant accident. Ultrasound Med Biol 1997; 23:9-13.
7. Warszawski A, Rottinger EM, Vogel R, Warszawski N. 20 MHz ultrasonic imaging for quantitative assessment and documentation of early and late

postradiation skin reactions in breast cancer patients. Radiother Oncol 1998; 47:241-7.
8. Thijssen JM. Ultrasonic tissue characterisation and echographic imaging. Phys Med Biol 1989; 34:1667-74.
9. Forster FK, Olerud JE, Riederer-Henderson MA, Holmes AW. Ultrasonic assessment of skin and surgical wounds utilizing backscatter acoustic techniques to estimate attenuation. Ultrasound Med Biol 1990; 16:43-53.
10. Raju BI, Swindells KJ, Gonzalez S, Srinivasan MA. Quantitative ultrasonic methods for characterization of skin lesions in vivo. Ultrasound Med Biol 2003; 29:825-38.
11. Huang YP, Zheng YP, Leung SF. Quasilinear viscoelastic parameters of neck tissues with fibrosis induced by radiotherapy. Clin Biomech 2005; 20: 145-154.
12. Leung SF, Zheng YP, Choi CYK, Mak SSS, Chiu SKW, Zee B, Mak AFT. Quantitative measurement of post-irradiation neck fibrosis based on Young's modulus: description of a new method and clinical results. Cancer 2002; 95: 656-662.
13. Zheng YP, Leung SF, Mak AFT. Assessment of neck tissue fibrosis using an ultrasound palpation system: A feasibility study. Med Biol Eng Comp 2000; 38: 1-6.
14. Guittet C, Ossant F, Remenieras JP, Pourcelot L, Berson M. High-frequency estimation of the ultrasonic attenuation coefficient slope obtained in human skin: simulation and in vivo results. Ultrasound Med Biol 1999; 25:421-9.
15. Fournier C, Bridal SL, Berger G, Laugier P. Reproducibility of skin characterization with backscattered spectra (12--25 MHz) in healthy subjects. Ultrasound Med Biol 2001; 27:603-10.

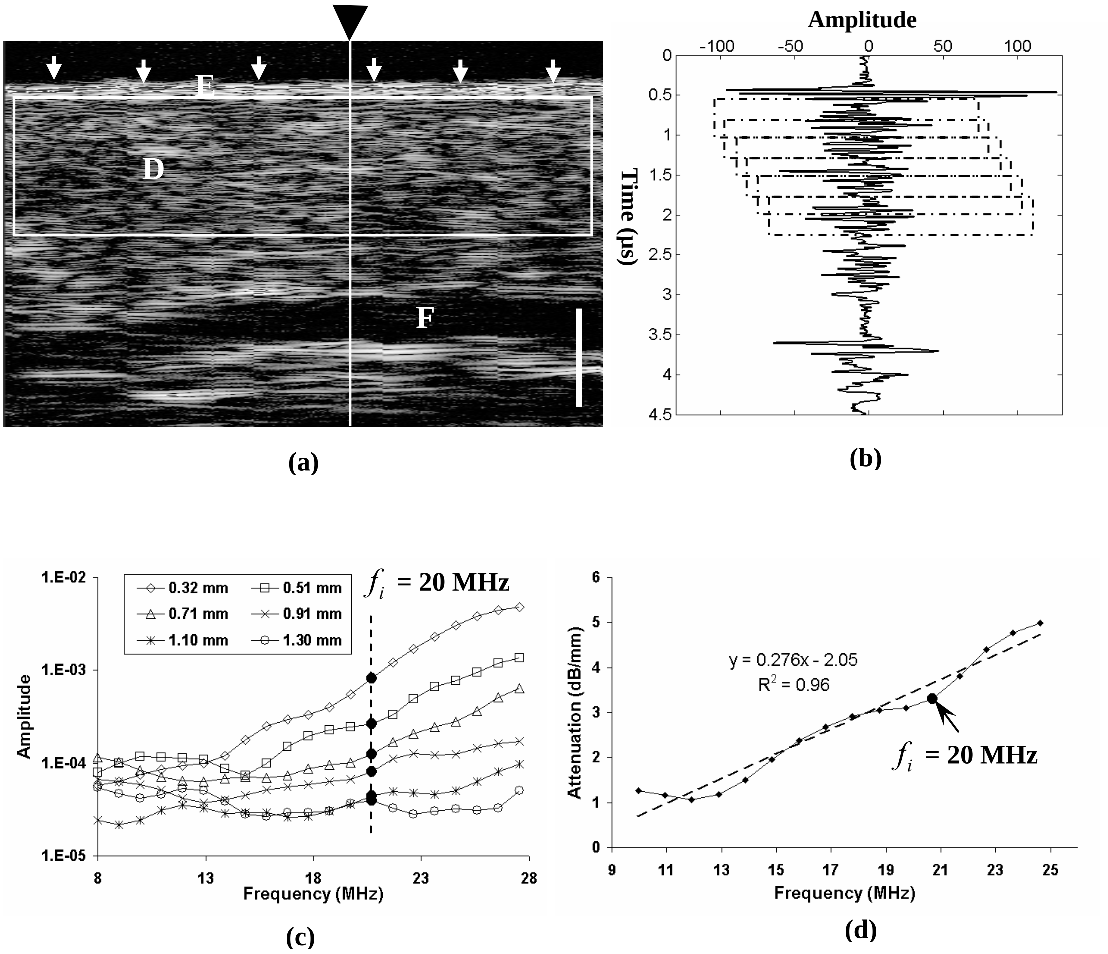
16. Serup J. Ten years' experience with high-frequency ultrasound examination of the skin: development and refinement of technique and equipments. In: Altmeyer P, el-Gammal S, Hoffmann K, eds. Ultrasound in dermatology. Berlin: Springer-verlag, 1992:41-54.
17. Bridal SL, Fornes P, Bruneval P, Berger G. Correlation of ultrasonic attenuation (30 to 50 MHz) and constituents of atherosclerotic plaque. Ultrasound Med Biol 1997; 23:691-703.
18. Raju BI, Srinivasan MA. High-frequency ultrasonic attenuation and backscatter coefficients of in vivo normal human dermis and subcutaneous fat. Ultrasound Med Biol 2001; 27:1543-56.
19. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clin Rehabil 1998; 12:187-99.
20. Lobnig BM, Bender R, Maslowska-Wessel E. Repeatability of heart rate variability measured via spectral analysis in healthy subjects. J Clin Basic Cardiol 2003; 6:29-33.
21. Pellacani G, Seidenari S. Variations in facial skin thickness and echogenicity with site and age. Acta Derm Venereol 1999; 79:366-9.

Figure captions:

Fig. 1. The processing of the RF signals for the extraction of the attenuation coefficients. (a) One typical RF image of the normal skin. The black arrow and the vertical line indicate where the RF signal in (b) locates. The white arrows indicate there is a spatial averaging of the power spectra along the lateral direction for calculating the ultrasonic properties. The white rectangle indicates the region of interest where the ultrasonic properties are calculated. The white bar in the lower right indicates a scale of 1 mm. **E** – skin entry and epidermis, **D** – dermis, **F** – subcutaneous fat. (b) One typical RF signal from the site where the black arrow and the vertical line indicate in (a). The dotted windows show where the power spectrum for each small region of interest is obtained. (c) The power spectra from the 6 small regions of interest which have been corrected for the system-dependent effects. At each frequency point as indicated by the dotted line, a regression of the power spectra with the axial distance is performed and then the attenuation coefficients are obtained. (d) The frequency-dependent attenuation coefficients of the dermis. The typical attenuation coefficient at 20 MHz is calculated where the dotted line in (c) indicates.

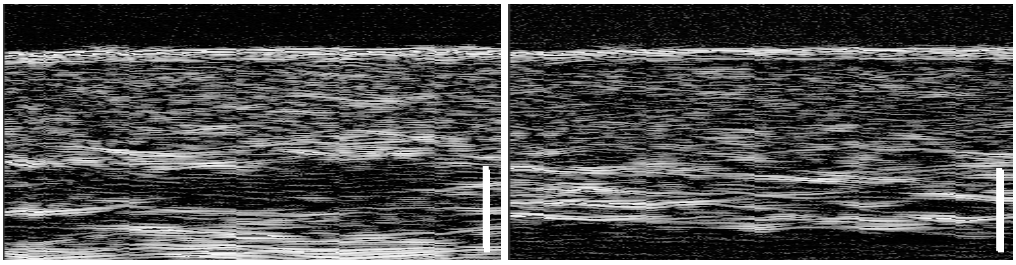
Fig. 2. Typical RF images of the skin observed in (a) forearm and (b) neck regions of a normal subject. The white bar in the lower right indicates a scale of 1 mm.

Fig. 3. Ultrasonic properties of the skin in the forearm and the neck regions of the normal subjects: (a) the attenuation slope β; (b) the integrated attenuation IA; (c) the integrated backscatter IBS.



E
D
F
(a)
Amplitude
-100
-50
0
50
100
Time (µs)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
(b)
0.32 mm
0.51 mm
0.71 mm
0.91 mm
1.10 mm
1.30 mm
f_i = 20 MHz
1.E-02
1.E-03
1.E-04
1.E-05
Amplitude
8
13
18
23
28
Frequency (MHz)
(c)
y = 0.276x - 2.05
R^2 = 0.96
f_i = 20 MHz
Attenuation (dB/mm)
0
1
2
3
4
5
6
9
11
13
15
17
19
21
23
25
Frequency (MHz)
(d)

Fig. 1.



(a) (b)

Fig. 2.

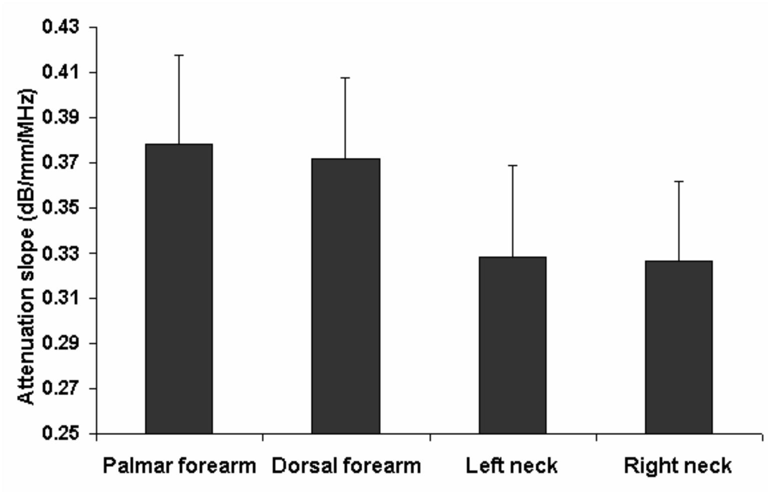


Fig. 3. (a)

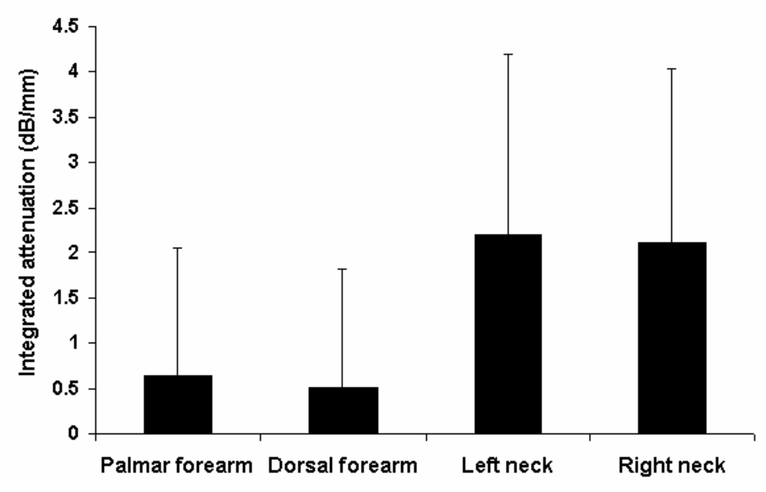


Fig. 3. (b)

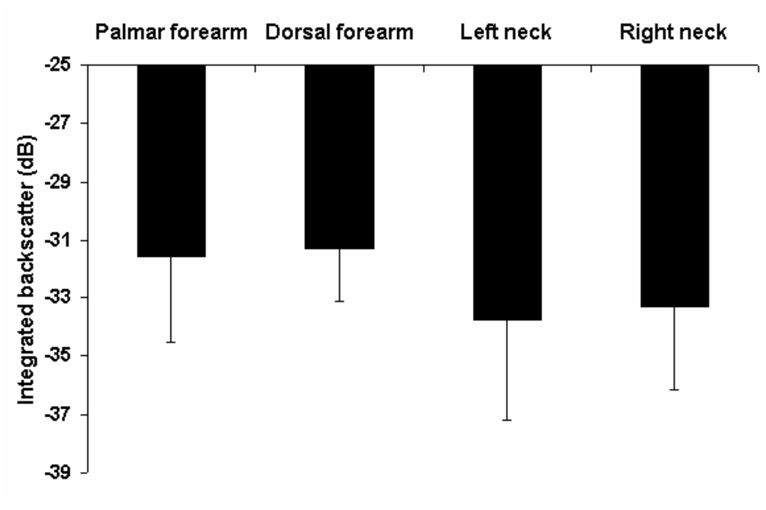


Fig. 3. (c)

TABLE 1. Intra- and inter-rater reliability of the ultrasonic (US) measurement

US parameters	***Intra-rater measurement***		***Inter-rater measurement***	
	ICC(1,3)	***95% CI***	***ICC(3,3)***	***95% CI***
β (dB/mm/MHz)	*0.87*	*0.75 - 0.93*	*0.83*	*0.67 - 0.91*
IA (dB/mm)	*0.86*	*0.73 - 0.92*	*0.91*	*0.83 - 0.95*
IBS (dB)	*0.88*	*0.77 - 0.93*	*0.97*	*0.94 - 0.98*

CI: confidence interval.

TABLE 2. (a) Bland-Altman test of the intra-rater measurement

US parameters	$\bar{d}$	$SD_{\bar{d}}$	***95% limits of agreement***	$SE_{\bar{d}}$	***95% CI of*** $\bar{d}$	***Reliability coefficient***
β (dB/mm/MHz)	*-0.001*	*0.031*	*-0.063 → 0.060*	*0.005*	*-0.011 → 0.009*	*0.063*
IA (dB/mm)	*0.09*	*0.95*	*-1.77 → 1.95*	*0.15*	*-0.21 → 0.39*	*1.91*
IBS (dB)	*-0.15*	*1.64*	*-3.37 → 3.07*	*0.26*	*-0.68 → 0.38*	*3.30*

(b) Bland-Altman test of the inter-rater measurement

US parameters	$\bar{d}$	$SD_{\bar{d}}$	***95% limits of agreement***	$SE_{\bar{d}}$	***95% CI of*** $\bar{d}$	***Reliability Coefficient***
β (dB/mm/MHz)	*0.004*	*0.035*	*-0.065 → 0.073*	*0.006*	*-0.007 → 0.015*	*0.071*
IA (dB/mm)	*-0.11*	*0.91*	*-1.89 → 1.68*	*0.14*	*-0.40 → 0.18*	*1.84*
IBS (dB)	*0.15*	*1.02*	*-1.84 → 2.15*	*0.16*	*-0.17 → 0.48*	*2.06*

Totally n = 40 measurements were used for computation of each test. $\bar{d}$ is the mean difference of the second measurement to the first measurement for the intra-rater operation, or the measurement of the second rater to that of the first rater for the inter-rater operation; $SD_{\bar{d}}$ is the SD of the mean difference calculated from the original data; $SE_{\bar{d}}$ is the standard error of the difference, calculated as $SE_{\bar{d}} = SD_{\bar{d}} / \sqrt{n}$; 95% CI of $\bar{d}$ is the 95% CI for the mean difference, being defined as $\bar{d} \pm t_{n-1} \cdot SE_{\bar{d}}$, where t is the critical value of *t*-distribution at $p = 0.05$. Reliability coefficient is calculated to be $2\sqrt{\sum_{i=1}^{n} d_i^2 / n}$, where d_i is the individual difference for test i among a total of n measurements.