Osteoarthritis and Cartilage 22 (2014) 1577–1582

Osteoarthritis and Cartilage

Brief Report

Articular cartilage degeneration classification by means of high-frequency ultrasound

N. Männicke †, M. Schöne †, M. Oelze ‡, K. Raum † *

† Julius Wolff Institute and Berlin-Brandenburg School for Regenerative Therapies, Charité-Universitätsmedizin Berlin, Germany
‡ Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

ARTICLE INFO

Article history:
Received 27 January 2014
Accepted 23 June 2014

Keywords:
Backscatter
Cartilage
Degeneration
Osteoarthrosis
Ultrasound biomicroscopy
Classification

SUMMARY

Context: To date only single ultrasound parameters were regarded in statistical analyses to characterize osteoarthritic changes in articular cartilage and the potential benefit of using parameter combinations for characterization remains unclear.
Objective: Therefore, the aim of this work was to utilize feature selection and classification of a Mankin subset score (i.e., cartilage surface and cell sub-scores) using ultrasound-based parameter pairs and investigate both classification accuracy and the sensitivity towards different degeneration stages.
Design: 40 punch biopsies of human cartilage were previously scanned *ex vivo* with a 40-MHz transducer. Ultrasound-based surface parameters, as well as backscatter and envelope statistics parameters were available. Logistic regression was performed with each unique US parameter pair as predictor and different degeneration stages as response variables. The best ultrasound-based parameter pair for each Mankin subset score value was assessed by highest classification accuracy and utilized in receiver operating characteristics (ROC) analysis.
Results: The classifications discriminating between early degenerations yielded area under the ROC curve (AUC) values of 0.94–0.99 (mean ± SD: 0.97 ± 0.03). In contrast, classifications among higher Mankin subset scores resulted in lower AUC values: 0.75–0.91 (mean ± SD: 0.84 ± 0.08). Variable sensitivities of the different ultrasound features were observed with respect to different degeneration stages.
Conclusions: Our results strongly suggest that combinations of high-frequency ultrasound-based parameters exhibit potential to characterize different, particularly very early, degeneration stages of hyaline cartilage. Variable sensitivities towards different degeneration stages suggest that a concurrent estimation of multiple ultrasound-based parameters is diagnostically valuable. *In-vivo* application of the present findings is conceivable in both minimally invasive arthroscopic ultrasound and high-frequency transcutaneous ultrasound.
© 2014 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Introduction

Ultrasound biomicroscopy (UBM) is capable of visualizing cartilage tissue at a high spatial resolution and gives access to a variety of quantitative parameters. Besides thickness, the most commonly derived quantitative parameters are surface reflection amplitude and surface roughness as surrogates for alterations of cartilage matrix stiffness and roughness, respectively. These parameters have been observed to significantly vary in the course of osteoarthrosis[1–4]. Moreover, the reflection intensity from the interface between cartilage and subchondral bone has been suggested to change due to a combination of increased sclerosis-related bone density and acoustic attenuation of the cartilage matrix[2].

Recently, we have shown that 3D-UBM not only enables improved estimation of surface properties[5], but also gives access to US backscatter parameters of the cartilage matrix[6], whose analyses have been only sparsely carried out until now. Statistically significant differences of individual surface and backscatter parameters were found with respect to early structural and cellular degenerations, as assessed by the histologically derived Mankin subset score. However, group differences were mostly observed between healthy samples (Mankin subset score 0) and all other samples having varying degrees of degeneration. Furthermore, a

* Address correspondence and reprint requests to: K. Raum, Julius Wolff Institute & Berlin-Brandenburg School for Regenerative Therapies, Charité – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. Tel: 49-30-450-539-503.
E-mail address: kay.raum@charite.de (K. Raum).

http://dx.doi.org/10.1016/j.joca.2014.06.019
1063-4584/© 2014 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

clear separation between different degeneration stages could not be obtained when using single parameters.

Therefore, the aim of this work was to combine the promising diagnostic potential of previously established surface reflection and matrix backscatter parameters by selecting relevant features with respect to different degeneration stages and employing them in classification and receiver operating characteristics (ROC) analyses. We hypothesized that ultrasound readings exhibit variable sensitivities with respect to different degeneration stages and that a combination of ultrasound parameters obtained from the cartilage surface and the sub-superficial tissue matrix will provide the ability to separate classes of degeneration, particularly between the early stages of cartilage degeneration.

Materials & methods

This work was based on the *ex-vivo* measurements, data evaluation and histological analysis of two previous studies[5,6]. The following three sections briefly summarize these aspects.

Samples

One to three punch biopsies (diameter: 8 mm) of cartilage were obtained from the femoral condyles of 19 patients during alloplastic implant surgery ($N = 38$). Two biopsies were excluded due to deep fissures or complete loss of cartilage. Moreover, one to three punch biopsies were obtained from the femoral joint of four human cadavers with no known degenerative joint disease ($N = 10$). In total, $N = 46$ biopsies were incorporated into the classification analysis.

After storage at −32°C, the biopsy specimens were immersed in Phosphate Buffered Saline (PBS) at 25°C and measured by UBM with the scanning acoustic microscope SAM200Ex[7,8]. A spherically focused 40-MHz transducer was used, providing a lateral and axial resolution of 120 and 50 μm, respectively. Samples were scanned in time-resolved C-scan mode, yielding one 3D dataset for every sample. The lateral scan-increments in both scan directions were 20 μm. Representative cross-sectional 2D images, and 3D fly-through videos are shown in the supplementary material.

Supplementary video related to this article can be found at http://dx.doi.org/10.1016/j.joca.2014.06.019.

Histology

Histological analysis was performed on demineralized and paraffin-embedded sections of the respective punch biopsies using routine histology processing and staining. Serial transverse sections (thickness: 5 μm) were cut through the central part of the biopsy. Cartilage degeneration was graded using the individual scoring categories (i.e., cartilage surface, cells, extracellular matrix proteins, and subchondral bone integrity) of the 14-point modified Mankin score[9]. The scoring was performed by two trained clinicians independently. When the scores were different, the scoring was revised[5]. Of particular interest in this study were the surface structure and the scoring of cellular abnormalities, denoted as M1 and M2, respectively. In the following, the Mankin subset score denotes the sum of M1 and M2. The Mankin subset scores of the 46 evaluated biopsies covered the following values: (0 [$n = 5$]; 1 [$n = 3$]; 2 [$n = 4$]; 3 [$n = 9$]; 4 [$n = 9$]; 5 [$n = 9$]; 6 [$n = 7$]). The excluded biopsies had the highest scores ($M1 = 6$; $M2 = 3$).

Parameter extraction

Ultrasound-based parameters were obtained from time gates originating from the cartilage surface (hereafter denoted surface parameters)[5] and from the cartilage matrix (backscatter parameters) at normal incidence regions[6]. In this work, nine ultrasound-based parameters were incorporated: At the surface, the spatial variation and the median value of the integrated reflection amplitude (*IRC*) yielded *ΔIRC* and *IRC* respectively and the temporal variability of the surface positions determined the Ultrasonic Roughness Index (*URI*)[5]. In six data sets, these parameters could not be calculated due to one of the following reasons: (1) surface region measured with small inclination (<5°) relative to the sound beam axis too small, (2) region of interest outside of focus range, (3) detached tissue fibers above cartilage surface. Depth-dependent profiles of backscatter amplitude (apparent integrated backscatter, *AIB*) and spectral slope (apparent frequency dependence of backscatter, *AFB*), were used to estimate the maximum values AIB_{max} and AFB_{max}, the depth-dependent slope AIB_{slope} and the extrapolation of the integrated backscatter to the cartilage surface AIB_0. Furthermore, in the transitional zone, backscattered waveforms were analyzed with envelope statistics, yielding k as ratio of coherent to incoherent signal energy and μ as scatterer number density per resolution cell[6,10].

Classification, feature selection, and ROC

To study the predictability with respect to different degeneration stages, the Mankin subset scores were divided into six binary classifications to distinguish between scores <i and ≥*i*, with $i = 1,2,3,4,5$ and 6.

Quasi-least squares (QLS) regressions[11] were used to account for the potential intra-individual correlation of biopsies obtained from the same donor. QLS were modeled using a Bernoulli-distributed outcome variable (i.e., the Mankin subset score discrimination) under the assumption of equicorrelated samples, i.e., all pairs of biopsies from one donor are expected to have the same correlation. Regression analyses were applied to all possible combinations of two ultrasound-based parameters as predictor variables and the six binary Mankin subset score discriminations as response variables. A binary operator (threshold: 0.5) was applied to the model output to facilitate binomial classifications. With leave-one-out cross-validation, the best ultrasound-based parameter pair for each Mankin subset score was assessed by means of highest classification accuracy. The latter was determined by the number of successful classifications divided by the total number of observations; a classification accuracy of 1 therefore denotes a perfect separation between the two classes. The classification scheme necessitates exclusion of samples for which not all parameters could be derived, thus only 40 samples were included. Due to the finite number of observations, several feature pairs could attain the highest classification accuracy. Therefore, ROC analysis was performed using QLS regression of the entire dataset without cross-validation and the area under the ROC curve (AUC) was calculated for all candidate pairs. The positive class label was assigned to the respective lower Mankin subset scores. The 95% confidence intervals were calculated by applying the bias corrected and accelerated percentile method with the use of 1000 bootstrap samples per analysis. Finally, for all six classifications, the ultrasound-based parameter pair with the highest AUC value was determined to be the best pair. All analyses were performed using custom-developed software based on the Statistics toolbox of Matlab (Matlab R2011b; Mathworks, Natick, MA, USA). QLS regression was performed using the GEEQBOX toolbox[12]. ROC analyses including the derivation of the AUC values were carried out using the "perfcurve" function from the Statistics Toolbox of Matlab.

Results

The highest cross-validated classification accuracies for the six classifications were in the range between 0.78 and 0.92 (mean ± SD:

Table I
Occurrence of ultrasound-based parameters in feature pairs exhibiting the highest classification accuracy of all unique feature pairs. One occurrence is indicated by one +. The feature pair with the highest AUC among all candidate feature pairs is indicated by gray cell background. The corresponding AUC values and 95% confidence intervals are listed below

		Mankin subset score discriminations					
		<1 | ≥1	<2 | ≥2	<3 | ≥3	<4 | ≥4	<5 | ≥5	<6 | ≥6
Surface	*IRC*	+	+	+			
	URI	+	+				+
	ΔIRC	+	+		+	+	+
Backscatter	*AIB0*		+		+		
	AIB_{slope}						
	AIB_{max}					++++	
	AFB_{max}	++++	+++	+		+	
Envelope statistics	*k*		+			+	
	μ	+				+	
Accuracy		0.925	0.925	0.900	0.775	0.825	0.875
AUC		0.99	0.99	0.94	0.75	0.87	0.91
(95% CI)		(0.92, 1.00)	(0.91, 1.00)	(0.82, 0.98)	(0.54, 0.89)	(0.70, 0.96)	(0.73, 0.98)

0.87 ± 0.06) (Table I). 4, 4, 1, 1, 4 and 1 potential feature pairs were found for the differentiation between Mankin subset scores smaller than 1, 2, 3, 4, 5, and 6, respectively. ROC analysis of the best feature pairs resulted in AUC values between 0.75 and 0.99 (mean ± SD: 0.91 ± 0.09). Classification accuracy and AUC values were highest and lowest for differentiating between Mankin subset scores up to three and between Mankin subset scores three and four, respectively. Among the individual ultrasound readings, the backscatter frequency-dependence parameter AFB_{max} in combination with surface and envelope statistics parameters were particularly predictive for differentiating between low Mankin subset scores (up to 3). In contrast, backscatter amplitude parameters AIB_{max} and AIB_0 were predominantly predictive for classifications between higher degeneration stages (>3). The surface and envelope statistics parameters were present in all Mankin subset score discrimination models. AIB_{slope} was not included in any pairwise combination with high classification accuracy.

Since a generalization of the selected features with respect to all degeneration stages is difficult, the classification accuracies of all feature candidate pairs were pooled for all six classifications. The highest average classification accuracy of 0.83 was achieved by the ultrasound parameter pair AIB_{max} and AFB_{max} with a standard deviation of 0.08.

Fig. 1 provides an overview of the class plots with the best ultrasound-based parameter pairs depicting the highest AUC values as well as the corresponding ROC curves for all six Mankin subset score classifications. Particularly, in the classification of low Mankin subset scores [Fig. 1(a–c)], a good class separation was observed. At full sensitivity, the specificities were 0.97, 0.94, and 0.79 for the separation of Mankin subset scores 0, smaller than 1, and smaller than 2, respectively. In contrast, the class separation between Mankin subset scores higher than 3 [Fig. 1(d–f)] was less apparent, resulting in lower AUC values for later stage cartilage degeneration.

Most parameters consistently separated early and advanced stages of degeneration as expected. Lower stages of degeneration were associated with high spectral slope values (AFB_{max}), small roughness values (*URI*), and low variations of the surface reflection values (*ΔIRC*), whereas in advanced degeneration stages, higher backscatter amplitude values (AIB_{max}) were observed. However, in the classification of Mankin subset scores smaller than 2, the class boundary suggests that lower *IRC* values are associated with earlier degenerations [Fig. 1(b)]. This observation can be attributed to two outliers at −30.7 dB and −30.6 dB with relatively high Mankin subset scores of 2 and 3, respectively. Moreover, the association of high *URI* values with earlier degenerations in the classification of Mankin subset scores smaller than 6 [Fig. 1(f)] is presumably an artifact due to the high class imbalance. Finally, because AFB_{max} alone yields a reasonable separation of Mankin subset score 0 from all other samples [Fig. 1(a)], the choice of scatterer number density per resolution cell *μ* as a second parameter should be interpreted with caution.

Discussion

This study follows up on the results of two previous studies and demonstrates that sophisticated ultrasound data analysis not only provides statistically significant differences of the derived parameters with respect to a histologically derived Mankin subset score, but elucidates the potential to use combinations of ultrasound-based parameters for the classification of progressive stages of degeneration.

One major challenge of this study was the relatively high ratio of feature number to number of observations, which can cause overfitting of the data[13]. Therefore, we decided to employ only feature pairs and select the best pair based on classification accuracy. Cross-validation was used to ascertain a selection of meaningful parameters. This feature selection yielded a description of sensitivity of the evaluated ultrasound parameters with respect to different stages of degeneration.

We observed that surface parameters and envelope statistics parameters are feature candidates throughout all degeneration stages, whereas backscatter amplitude and frequency dependence appear as good predictors for advanced and early stages of degeneration, respectively. Cell number density and backscatter amplitude are moderately correlated, suggesting that chondrocytes are an important scattering source in cartilage[6]. The feature selections strongly suggest that the frequency dependence of backscatter is sensitive to early cellular degenerative changes, e.g., chondron swelling, whereas backscatter amplitude decrease can be associated with hypocellularity. Interestingly, AIB_{slope} was not selected in any top pairwise combination although it has been found to be a highly predictive parameter for cartilage tissue repair, presumably due to its sensitivity to collagen packing density[8]. In line with previous studies[1–4,14], the classifications predicted higher *URI* and lower *IRC* values for advanced degeneration stages. Although no statistical difference with respect to different degeneration stages was present[5,6], surface parameters were selected as feature candidates throughout all degeneration stages, strengthening the observation that *URI* and *IRC* are subject to gradual changes in the progression of osteoarthrosis[2,3]. Feature pairs that were found most suitable for all Mankin subset score discriminations were backscatter amplitude in combination with spectral slope and surface roughness. This observation further highlights

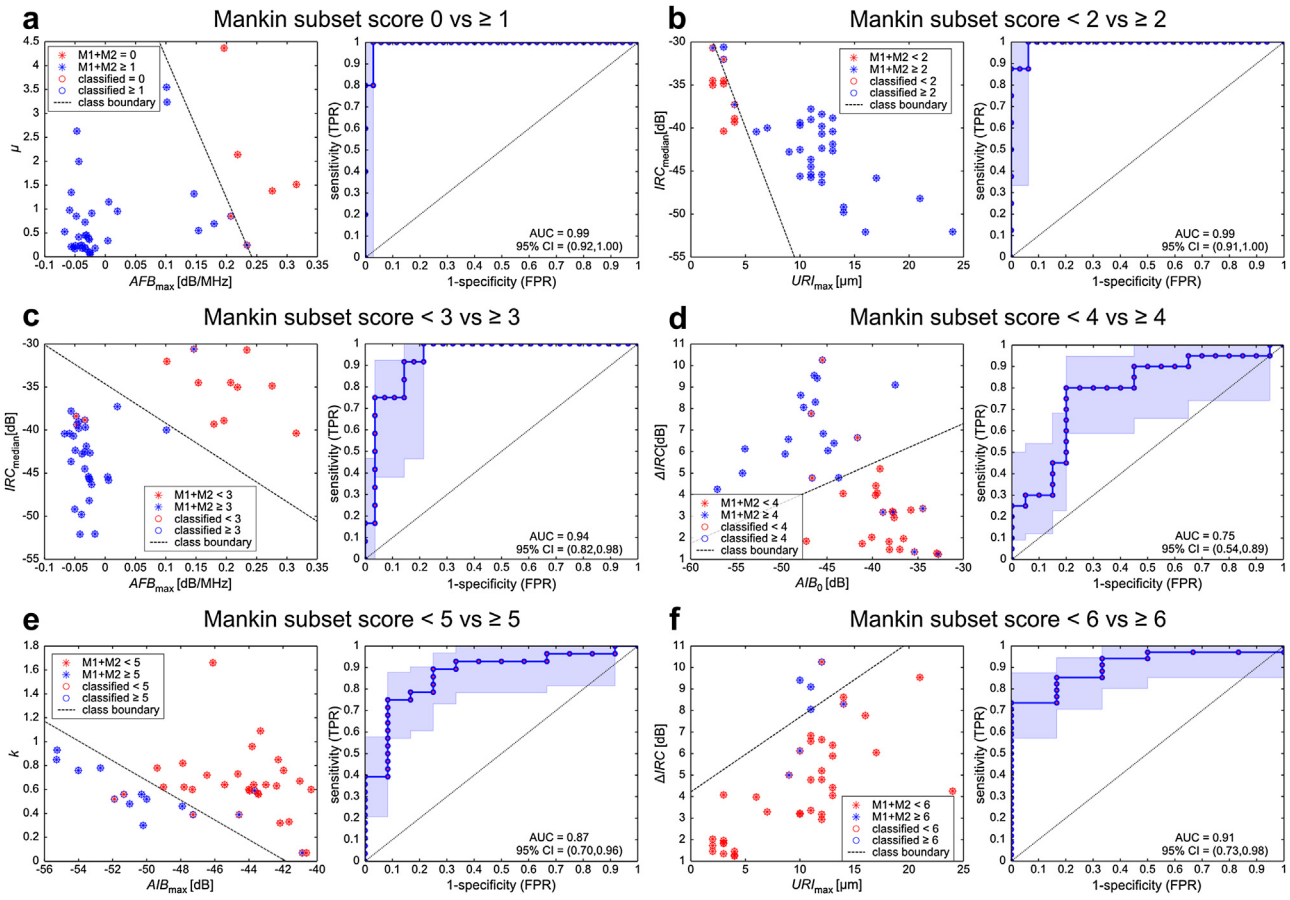

Fig. 1. Class plots (left panels) with the best performing ultrasound-based parameter pairs and the corresponding ROC curves (right panels) for all six Mankin subset score classifications. The positive (red data points) and negative (blue) class labels represent lower and higher Mankin scores, respectively. Classification output is highlighted by circles in blue and red, the class boundaries are indicated by black dashed lines. The right panels display the ROC curves in blue. The curves with no predictive value are illustrated by dotted lines.

the diagnostic value of ultrasound backscatter parameters combined with well-established surface parameters.

Note that due to the unbalanced class distribution for very early and advanced degeneration stages, these classification accuracies are possibly biased and should be interpreted with caution. However, ROC analysis is known to provide valid results also in the presence of unbalanced classes[15]. A further limitation was that 2D histology sections used for the histological grading represented only a sub-region of the 3D volumes evaluated by ultrasound. Since local changes are likely, particularly in degenerated tissue, the histological scores cannot be expected to be fully comparable to the ultrasound-based parameters. The samples from two donors with the highest scores could not be included in the ultrasound-based analysis due to the pronounced loss of cartilage matrix. In these samples, the high degree of degeneration was clearly visible in the acoustic images (not shown). For the other six excluded data sets, the surface parameters could not be calculated, either because the measured sample volume did not contain a sufficiently large region with a sufficiently small surface tilt and the required distance transducer and sample, or because detached tissue fibers prevented the correct surface reconstruction using our custom-developed algorithms[5]. While the former limitations can be overcome by using ultrasound array transducers, allowing multi-angle and multi-focus acquisitions, the latter requires further development of the surface reconstruction algorithms in the presence of detached fibers. Nevertheless, despite these limitations, we have shown for the first time that a combination of ultrasound-based parameters exhibit potential to characterize different, particularly the earliest, degeneration stages of human hyaline cartilage with AUC values up to 0.99.

In-vivo application of high-frequency ultrasound for OA assessment can be achieved in various ways. Transcutaneous ultrasound with high frequencies can access reasonably large cartilage areas at the femur condyles[16]. Moreover, minimally invasive arthroscopic ultrasound has been proposed as an alternative to an intra-operative application and ultrasound backscatter amplitude and surface parameters were successfully derived[14]. An extension to envelope statistics and spectral slope parameters is straightforward; however, accuracies and robustness must be elaborated for the respective ultrasonic transducers.

The most advanced technology for *in-vivo* grading of osteoarthritis in cartilage is magnetic resonance imaging (MRI). The whole-organ MRI score (WORMS) has been established to score the structural integrity of the entire joint including articular cartilage integrity[17]. With respect to the cartilage tissue, this and other morphological scoring systems focus on the reduction of cartilage thickness and the occurrence and extend of focal defects, which are also clearly visible in high-frequency ultrasound images (see Supplementary Figure). Special MRI sequences have been proposed to grade OA related biochemical changes. Delayed gadolinium-enhanced (dGEMRIC) T_1 imaging has been shown to be sensitive to changes of the proteoglycan (PG) content in several studies[18,19]. Other relaxation times, e.g., $T_{1\rho}$ and T2 have also been

proposed as good overall indicators of cartilage health, as described by the Mankin score[20]. Although significant correlations between T2* values with total Mankin score ($R = -0.362$) and the cartilage surface morphology subscore ($R = -0.367$) have been observed in 21 human femoral head specimens with varying severities of OA by Bittersohl *et al.*[21], the specific contrast mechanisms and their sensitivities with respect to specific tissue alterations other than changes of the PG content have yet to be established.

In conclusion, the combination of ultrasound-based parameters derived from surface reflections and signals backscattered from the cartilage matrix provides a promising capability to distinguish between different cartilage degeneration stages, particularly between the earliest stages. The surface parameters used in this study are related to cartilage softening and surface fibrillation[5,8,22], while the backscatter parameters have been shown to be associated with changes in cell morphology, i.e., cell density and clustering[6]. Such changes usually occur prior to a destruction of the cartilage matrix.

However, ultrasound-based parameters have variable sensitivities towards different degeneration stages, suggesting that a concurrent estimation is diagnostically valuable. Future work will also need to unravel the structural constituents that contribute to ultrasound backscatter, as a better understanding of the backscatter mechanisms is anticipated to further improve the ultrasound-based classification of OA. Moreover, a direct comparison between ultrasound and MRI based scores is needed to elucidate the diagnostic value and potential limitations of the US-based classification approach.

Contributions

- Nils Männicke: (1) Acquisition of data, Analysis and interpretation of the data, (2) Drafting of the article, (3) Final approval of the article
- Martin Schöne (1) Acquisition of data, (2) Critical revision of the article for important intellectual content, (3) Final approval of the article
- Michael L Oelze (1) Conception and design, (2) Critical revision of the article for important intellectual content, (3) Final approval of the article
- Kay Raum (1) Acquisition of data, Conception and design, (2) Drafting of the article, (3) Final approval of the article

Role of funding

This work was supported by the German Research Council (DFG grant Ra1380/6) and Deutsche Arthrose-Hilfe e.V. N.M. received a DFG stipend through the Berlin-Brandenburg School for Regenerative Therapies GSC 203.

All funding sources were not involved in the study design, in the collection, analysis and interpretation of data or in the preparation of the manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank Dr Andreas Busjahn for his advice regarding the statistical analysis.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.joca.2014.06.019.

References

1. Nieminen HJ, Zheng Y, Saarakkala S, Wang Q, Toyras J, Huang Y, *et al.* Quantitative assessment of articular cartilage using high-frequency ultrasound: research findings and diagnostic prospects. Crit Rev Biomed Eng 2009;37:461–94.
2. Saarakkala S, Wang SZ, Huang YP, Jurvelin JS, Zheng YP. Characterization of center frequency and bandwidth of broadband ultrasound reflected by the articular cartilage to subchondral bone interface. Ultrasound Med Biol 2011;37:112–21.
3. Kaleva E, Saarakkala S, Toyras J, Nieminen HJ, Jurvelin JS. In-vitro comparison of time-domain, frequency-domain and wavelet ultrasound parameters in diagnostics of cartilage degeneration. Ultrasound Med Biol 2008;34:155–9.
4. Wang SZ, Huang YP, Saarakkala S, Zheng YP. Quantitative assessment of articular cartilage with morphologic, acoustic and mechanical properties obtained using high-frequency ultrasound. Ultrasound Med Biol 2010;36:512–27.
5. Schöne M, Männicke N, Gottwald M, Göbel F, Raum K. 3-D high frequency ultrasound improves the estimation of surface properties in degenerated cartilage. Ultrasound Med Biol 2013;39:834–44.
6. Männicke N, Schöne M, Gottwald M, Göbel F, Oelze ML, Raum K. 3-D high-frequency ultrasound backscatter analysis of human articular cartilage. Ultrasound Med Biol 2014;40:244–57.
7. Petter-Puchner A, Gruber-Blum S, Walder N, Fortelny RH, Redl H, Raum K. Ultrasound biomicroscopy (UBM) and scanning acoustic microscopy (SAM) for the assessment of hernia mesh integration: a comparison to standard histology in an experimental model. Hernia 2013, http://dx.doi.org/10.1007/s10029-013-1201-9.
8. Gelse K, Olk A, Eichhorn S, Swoboda B, Schoene M, Raum K. Quantitative ultrasound biomicroscopy for the analysis of healthy and repair cartilage tissue. Eur Cell Mater 2010;19:58–71.
9. van der Sluijs JA, Geesink RG, van der Linden AJ, Bulstra SK, Kuyer R, Drukker J. The reliability of the Mankin score for osteoarthritis. J Orthop Res 1992;10:58–61.
10. Hruska DP, Oelze ML. Improved parameter estimates based on the homodyned K distribution. IEEE Trans Ultrason Ferroelectr Freq Control 2009;56:2471–81.
11. Shults Justine, Hilbe Joseph M. Quasi-least Squares Regression. 1st edn. CRC Press; 2014.
12. Ratcliffe SJ, Shults J. GEEQBOX: a MATLAB toolbox for generalized estimating equations and quasi-least squares. J Statistical Softw 2008;25:1–14.
13. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 2004;66:411–21.
14. Liukkonen J, Hirvasniemi J, Joukainen A, Penttila P, Viren T, Saarakkala S, *et al.* Arthroscopic ultrasound technique for simultaneous quantitative assessment of articular cartilage and subchondral bone: an in vitro and in vivo feasibility study. Ultrasound Med Biol 2013;39:1460–8.
15. Fawcett T. An introduction to ROC analysis. Pattern Recogn Lett 2006;27:861–74.
16. Grassi W, Lamanna G, Farina A, Cervini C. Sonographic imaging of normal and osteoarthritic cartilage. Semin Arthritis Rheum 1999;28:398–403.
17. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, *et al.* Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage 2004;12:177–90.
18. Zilkens C, Miese F, Herten M, Kurzidem S, Jager M, Konig D, *et al.* Validity of gradient-echo three-dimensional delayed gadolinium-enhanced magnetic resonance imaging of hip

joint cartilage: a histologically controlled study. Eur J Radiol 2013;82:e81–6.
19. Crema MD, Hunter DJ, Burstein D, Roemer FW, Li L, Eckstein F, *et al*. Association of changes in delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) with changes in cartilage thickness in the medial tibiofemoral compartment of the knee: a 2 year follow-up study using 3.0 T MRI. Ann Rheum Dis July 2013 [Epub ahead of print].
20. Li X, Cheng J, Lin K, Saadat E, Bolbos RI, Jobke B, *et al*. Quantitative MRI using T1rho and T2 in human osteoarthritic cartilage specimens: correlation with biochemical measurements and histology. Magn Reson Imaging 2011;29:324–34.
21. Bittersohl B, Miese FR, Hosalkar HS, Herten M, Antoch G, Krauspe R, *et al*. T2* mapping of hip joint cartilage in various histological grades of degeneration. Osteoarthritis Cartilage 2012;20:653–60.
22. Kaleva E, Saarakkala S, Jurvelin JS, Viren T, Toyras J. Effects of ultrasound beam angle and surface roughness on the quantitative ultrasound parameters of articular cartilage. Ultrasound Med Biol 2009;35:1344–51.