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.

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/

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: