



Published in final edited form as:

IEEE Sens J. 2023 October 23; 23(23): 29619–21629. doi:10.1109/jsen.2023.3325153.

Classifying Pre-Radiographic Osteoarthritis of the Knee Using Wearable Acoustics Sensing at the Point of Care

Christopher J. Nichols,

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

Göktu C. Özmen [Graduate Student Member, IEEE],

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

Kristine Richardson [Graduate Student Member, IEEE],

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

Omer T. Inan [Senior Member, IEEE],

School of Electrical and Computer Engineering and by courtesy, the Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology.

Dave Ewart, MD

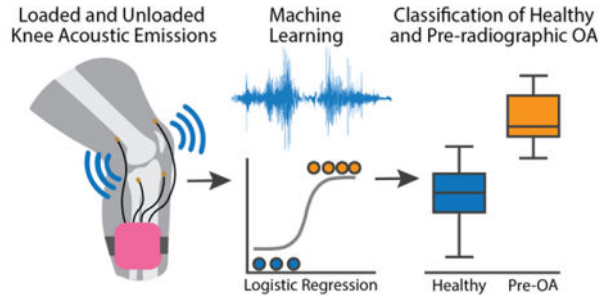
Minneapolis Veterans Affairs Medical Center, Minneapolis, MN.

Abstract

This study was undertaken to determine if knee acoustic emissions (KAE) measured at the point of care with a wearable device can classify knees with pre-radiographic osteoarthritis (pre-OA) from healthy knees. We performed a single-center cross-sectional observational study comparing KAE in healthy knees to knees with clinical symptoms compatible with knee OA that did not meet classification criteria for radiographic knee OA. KAE were measured during scripted maneuvers performed in clinic exam rooms or similarly noisy medical center locations in healthy (n=20), pre-OA (n=11), and, for comparison, OA (n=12) knees. Acoustic features were extracted from the KAE and used to train models to classify pre-OA, OA, and control knees with logistic regression. Model performance was measured and optimized with Leave-One-Out Cross-Validation. Regressive sensitivity analysis was performed to combine acoustic information from individual maneuvers to further optimize performance. Test-retest reliability of KAE was measured with intraclass correlation analysis. Classification models trained with KAE were accurate for both pre-OA and OA (94% accurate, 0.96 and 0.99 area under a receiver operating characteristic curve (AUC), respectively). Acoustic features selected for use in the optimized models had high test-retest reliability by intrasession and intersession intraclass correlation analysis (mean intraclass correlation coefficient 0.971 +/- 0.08 standard deviation). Analysis of KAE measured in acoustically uncontrolled medical settings using an easily accessible

wearable device accurately classified pre-OA knees from healthy control knees in our small cohort. Accessible methods of identifying pre-OA could enable regular joint health monitoring and improve OA treatment and rehabilitation outcomes.

Graphical Abstract



Index Terms—

pre-radiographic osteoarthritis; early osteoarthritis; knee; acoustics; crepitus

I. Introduction

Osteoarthritis of the knee is typically diagnosed at advanced stages when patients present with persistent knee pain and radiographic changes on x-ray. Radiographic knee OA is defined as Kellgren-Lawrence (KL) grade 2 or higher changes such as osteophyte formation and joint space narrowing [1]. Knee OA is a slowly progressive disease and by the time a knee has KL2 changes, the disease has been progressing for many years and the joint has considerable mechanical dysfunction that may hinder the efficacy of non-surgical therapies [2]. As such, efforts to develop disease-modifying therapies at this stage of advanced radiographic osteoarthritis (OA) have been unsuccessful [3], [4]. There is considerable interest in defining, characterizing, and identifying an earlier stage of knee OA, such as pre-radiographic OA (pre-OA) where an individual has stereotypic symptoms without such advanced structural changes, that may be more amenable to treatment or prevention [5]. Clinically available tools for diagnosis of pre-OA are insufficient—x-rays lack sensitivity and MRI is overly sensitive but lacking specificity for symptoms of disease or risk of progression to OA [6]–[9]. Further, x-rays require radiation exposure and MRI is time-consuming and expensive. Better methods for identification of pre-OA are needed.

The human knee generates noise and vibration which can be heard and felt, termed crepitus, that has long been associated with and is part of some clinical classification criteria for knee OA [10]. Evidence for crepitus as a sign of an earlier knee OA disease state is contradictory. In the Rotterdam Study, crepitus either heard or palpated by study personnel was associated with incident patellofemoral, but not tibiofemoral OA changes on MRI [11]. In the Osteoarthritis Initiative (OAI), subjective report of crepitus on the Knee Injury and Osteoarthritis Outcome Score (KOOS) was associated with incident symptomatic OA primarily in participants with preexisting asymptomatic tibiofemoral radiographic OA, while

palpable vibratory crepitus was not associated with incident radiographic OA development over four years in participants with baseline frequent knee symptoms [12], [13]. Crepitus is an inherently qualitative and unreliable metric and thus, it is unlikely that the clinically observed phenomenon of crepitus is a sensitive or specific enough measure to identify individuals with early disease [14].

Knee acoustics, termed knee acoustic emissions (KAE), are a highly sensitive and quantifiable recording of crepitus measured with airborne or contact microphones (accelerometers). KAE can be recorded non-invasively with a wearable device in acoustically uncontrolled settings and during dynamic movements, as well as in all knees, whether healthy or pathologic, for comparison in the context of knee health status [15]–[17]. Further, KAE contain information about the biomechanical and tribological forces experienced by the knee and can differentiate knees with radiographic OA from knees without OA [18]–[22]. To determine if quantitative analysis of crepitus via KAE has the potential to identify or characterize early knee OA, we undertook a pilot study first measuring knee acoustics in clinical settings using a wearable device and then combining them with machine learning techniques to predict whether the recording was from a knee with pre-OA or from a healthy knee.

II. Methods

A. Participant Recruitment Criteria

Participants were recruited from the Minneapolis Veterans Affairs Medical Center (MVAMC). This study was approved by the MVAMC Institutional Review Board and participants signed written informed consent (IRB #1594824). Pre-OA was defined as a participant having persistent knee pain but not meeting classification criteria for radiographic knee OA. Persistent knee pain was defined as an affirmative answer to the question “Have you had pain within the last year in or around the knee that occurred on most days for at least a month?” [20]. Radiographic knee OA classification criteria, which were also used to define our OA participant group, are defined as grade II or higher on the standardized KL grading scale used to assess the severity of knee OA [23]. Pre-OA participants with KL 0 x-rays had an MRI of the knee demonstrating damage to the tibiofemoral articular cartilage [24]. X-rays and MRIs were read by a trained musculoskeletal radiologist and reviewed by the primary author (DE). Healthy knees were defined as those without symptoms by patient self-report and no history of serious knee trauma or surgery. In a subset of participants with persistent knee pain, symptoms were further characterized with the Knee Injury and Osteoarthritis Score (KOOS) [25]. Participants who answered anything other than “never” for KOOS symptoms question 2 (S2: “Do you feel grinding, hear clicking or any other type of noise when your right knee moves?”), were designated as having crepitus [12]. Patients who had active inflammatory arthritis, or who had undergone knee surgery or had serious knee trauma in the past year were excluded.

B. Wearable System Design

Our group designed a fully wearable device, that is made of relatively inexpensive components and can be assembled outside of a niche research laboratory, for quickly measuring KAEs with high-fidelity in clinical environments. The device consists of a printed circuit board previously developed by our group, 3-D printed case, 500 mAh lithium-ion battery, and four piezoelectric contact microphones (BU-23173-000, Knowles, Itasca, IL) (Fig. 1A) [15]. Schematics and detailed written instructions were used to produce new devices by an independent team at our clinical recording site (MVAMC). To summarize elements of our system, the printed circuit board used a custom analog front-end, a four-channel analog-to-digital (ADC) converter, and the SAM4L8 microcontroller (Microchip Technology Inc., Chandler, AZ, USA) to record KAEs simultaneously from multiple locations around the knee joint [15]. Contact microphones were selected for their compact design, large bandwidth, and robustness to ambient noise artifacts, and have been successful in capturing KAEs in existing literature [26]–[28]. Microphones were sampled at 46.875 kHz, with microphone characterization showing linear response until approximately 12 kHz [29]. Our analyses filtered all data to at most below 5 kHz where the vast majority of KAE spectral energy is located, thereby being unaffected by this non-linearity [30]. All data were saved locally to a 32-gigabyte micro-SD card. A micro-USB port was used both for downloading saved data from the micro-SD card as well as battery charging [15]. Untethered design, large storage capacity and extended battery life allowed for several hours of recording as well as rapid donning and doffing of the system. The cost of our device's components was approximately \$725.

Our contact microphones were positioned medial and lateral to the superior and inferior aspects of the patella using double sided adhesive tape dots (3M, St. Paul, MN). Microphones were placed at the pre-specified locations anatomically by study personnel with clinical training in the evaluation and treatment of knee disorders who had expert knowledge of knee anatomy. Microphone cables were highly flexible and attached to the leg via an elastic band to minimize the potential for cable bumping noise artifacts [31]. The main housing was secured to each patient's shank using a soft Velcro strap to allow for flexible sizing. This configuration was used for all measurements (Fig. 1B).

C. Measurements of Knee Acoustic Emissions

Participants were asked to perform scripted maneuvers, consisting of flexion and extension, sit-to-stand, and walking, while recording KAE for 40 seconds (Fig. 1C and Fig. 1D) in clinic exam rooms or office locations in the medical center with similarly uncontrolled ambient noise. Knee angle was measured with a dual-axis goniometer (Biometrics, Newport, UK) for flexion-extension. Participants were encouraged to perform flexion-extension and sit-to-stand at a standardized pace (0.25 Hz) using a visual timer and to walk at a self-selected comfortable pace. To preserve data, if a participant was unable to complete a scripted maneuver or there was an error in acoustic measurement identified after the participant's visit, data for only that maneuver was excluded from the dataset.

D. KAE Preprocessing and Feature Extraction

KAE were filtered from 250 Hz to 5 kHz for flexion-extension and sit-to-stand, and 150 Hz to 1 kHz for walking (Fig. 1E) [19]. All maneuvers were then segmented into movement cycles: flexion-extension using goniometer data, sit-to-stand and walking data using energy peaks in the acoustic signal envelope occurring from the impact of sitting onto an exam table and heel strike, respectively. Flexion-extension and sit-to-stand cycles faster than 0.4 Hz were excluded to standardize joint loading conditions [18], [32]. No cycle exclusion criteria were used for walking. Cycles were subdivided into 30 ms windows with 20 ms overlap, and 59 spectral, temporal, Mel-cepstral, and band power-derived acoustic features were extracted from each window. These features were selected to characterize the pace, shape, and amplitude of the acoustic signal and have been previously validated for assessing joint health in other articular diseases [18], [26], [33]. Windowed features were then averaged across each cycle to improve signal robustness to outliers and quantify feature distribution within cycles. A feature matrix of m rows of movement cycles by n columns of features was created for each microphone, as well as a combined matrix which concatenated features from all microphones. Cycles were labeled one of three classes: healthy, pre-OA, or OA. This process was repeated for all maneuver types. All KAE preprocessing, feature extraction, and feature visualizations were performed using MATLAB (The MathWorks Inc. Natick, MA).

E. KAE Data Analysis

1) Comparison of Acoustics Between Healthy, Pre-OA, and OA knees—

Acoustic Features were averaged across movement cycles and compared between groups using density plots to estimate feature distribution probabilities for each group. After finding non-normal feature distributions with the Shapiro-Wilk Test, Kruskal-Wallis Test was used to compare feature medians between groups. Dunn's test was used for post-hoc analyses with the Bonferroni correction using RStudio (PBC, Boston, MA). Principal component analysis (PCA) was also used to reduce the dimensionality of all acoustic features to further visualize data variability. Movement cycles were color-labeled by disease status of the knee from which they were measured, and their first two principal components were plotted for aggregate feature comparison.

2) Model Development for Classifying Knee Health Status—

We used acoustic features to train classification models for predicting knee health status using binomial and multinomial logistic regression and measured model performance with Leave-One-Out Cross-Validation (Fig. 1E). All machine learning algorithms and visualizations were performed in Python (Scotts Valley, CA). Acoustic features were organized into feature sets by scripted maneuver movement cycle and health status label (healthy, pre-OA, OA). Feature sets were separated into training and testing sets, the training set being all knees except the "left-out" knee, which was the testing set. For participants with bilateral measurements, data from the participant's contralateral knee was excluded from the training set when testing his or her index knee to avoid data leakage. Training and testing sets were then converted to a standard normal distribution of the training data with a mean of zero and standard deviation of one. Training set knee acoustics were used to train each model, which was then tested on the left-out knee. This cross-validation was repeated for all knees

for each scripted maneuver. Recursive feature elimination was used during each round of cross-validation on the training data to reduce model inputs to acoustic features which best optimized performance by scripted maneuver and disease state.

Two primary models were developed to compare classification of pre-OA with KAEs compared to that of OA: healthy vs pre-OA, and healthy vs OA. Two additional models were also evaluated using all three knee health status classes: healthy vs all grades of OA, and multi-class classification of the three groups. Each model estimated the probabilities for acoustic features from a given movement cycle being produced by each respective health status classification. For our binary models, this predicted whether KAEs were produced by a “healthy” knee or an “injured” knee, while our multi-class classifier estimated the likelihood KAEs were produced by a “healthy”, “pre-OA”, or “OA” knee. Probabilities for each health status class were averaged across all cycles for each scripted maneuver to produce a ‘Knee Score’ for each maneuver. A composite Knee Score was calculated by combining and weighting Knee Scores from each individual scripted maneuver. We used regressive sensitivity analysis to determine the optimal weighting of each scripted maneuver’s Knee Score with a second stage of classification using Knee Scores as input features for logistic regression to optimize accuracy measured with Leave-One-Out Cross-Validation. Missing Knee scores were imputed from the average of a participant’s other scores for any who did not have data for all three maneuvers. By rounding each participant’s Knee Score to the closest number – 0 for “healthy”, 1 for “injured” – to determine disease status predictions, the accuracy, sensitivity, and specificity each model was calculated for each individual maneuver as well as the combination of maneuvers. Overall model performance of binary classifiers was quantified by calculating the area under a receiver operating characteristic curve (AUC). Because the DeLong test, a commonly reported methodology for comparison of AUC, is not appropriate for nested models derived from the same dataset, composite model performances were compared to that of a model using only age and BMI as features by bootstrapping Knee Score distributions to calculate the means, standard deviations, and 95% confidence intervals for the difference in each model’s mean AUC [34], [35]. Mean and standard deviation of Knee Scores aggregated for all three scripted maneuvers were compared between “healthy” and “injured” labeled knees with the healthy vs pre-OA and healthy vs OA models using Mann-Whitney U tests.

3) Evaluation of KAE Test-Retest Reliability—The reliability of KAE was evaluated for 10 knees of varying disease severity from five participants. Participants’ KAE were measured four times across two measurement days. The first measurement was captured from both knees across all three maneuvers. The wearable system was then completely removed, the electronics swapped between knees, and all maneuvers were recorded a second time. Participants returned one week later to repeat this process for the third and fourth measurements. Acoustic features were then extracted from all four measurements in the same manner as used in our classification analysis for each scripted maneuver. Intra-class correlation analysis as used to compare intrasession and intersession reliability of features recorded from each respective microphone measurement location. Features with an intraclass correlation coefficient (ICC) value below 0.75 were deemed to have poor repeatability and were therefore excluded from model feature selection [36].

III. Results

A. Participants

KAE were measured from 43 knees (20 control, 12 OA, 11 pre-OA) from 32 participants. Mean age of participants was 52.2 years (SD 18.5 years) and mean BMI was 31.6 (SD 7.9). 71% of participants were male. Five pre-OA knees had KL 0 with tibiofemoral articular cartilage damage on MRI and the other six had KL 1. Of the 5 KL0 knees, 4 had degenerative medial meniscus tears; none had lateral meniscus tears; 4 of 5 had patellofemoral cartilage damage; 4 had medial tibiofemoral cartilage damage, and the other had lateral tibiofemoral cartilage damage. Five radiographic OA knees were KL2, 5 were KL3, and 2 were KL4. Demographics and per group measurements by scripted maneuver are included in Supplementary Table 1 - 2, respectively. KOOS subscales and frequency of crepitus for the pre-OA and OA participants who completed a KOOS questionnaire are included in Supplementary Table 3.

B. Comparison of Acoustic Features

Groupwise distributions of acoustic features, selected for model performance optimization during cross-validation, demonstrated changes in feature densities from healthy, to pre-OA, and to OA (Fig. 3). Some mean feature values were significantly different between groups, particularly for flexion-extension and sit-to-stand feature sets (Fig. 3, Supplementary Table 4). PCA of aggregate acoustic features generally showed a separation between healthy knees and knees with OA, with pre-OA knees clustered between (Supplementary Fig. 1). Cluster separation of groups was most evident during flexion extension and sit-to-stand and from the inferomedial and inferolateral microphones.

C. Acoustic Feature Selection

Acoustic features selected for model performance optimization were different between scripted maneuvers and, to a lesser extent, model (Fig. 4). Flexion-extension features chosen for classification of pre-OA, such as spectral centroid and the first Mel-frequency Cepstral Coefficient (MFCC 1), evaluated the shape of the acoustic spectrum and emphasized the larger concentration of signal power at lower frequencies in arthritic knees, while OA classification favored power-based and cepstral features for optimal performance [26]. Both sit-to-stand models emphasized cepstral features and zero-crossing rate, which is often used for identifying high-pitched sounds from percussive noise [37]. Features selected for walking more broadly represented the acoustic energy spectrum and included many spectral features.

D. Classification Performance of Pre-Radiographic Knee OA

Mean composite Knee Scores, the models' probability estimate of a knee being "arthritic", for healthy knees were lower than for pre-OA (0.30 \pm 0.20 vs 0.81 \pm 0.15, $p < 0.001$, 95% confidence interval for between-group difference (CI) : 0.37 – 0.65) and for OA (0.24 \pm 0.22 vs 0.82 \pm 0.13, $p < 0.001$, CI: 0.45 – 0.71) (Fig. 5A and 5B, Table 1 and Supplementary Table 5, respectively). Accuracy of our acoustic classifier trained with the combined dataset from all three scripted maneuvers was higher than classifiers trained with

any individual maneuver's dataset (Fig. 6). As such, we chose the combined maneuver acoustic classifiers as our model of choice and compared its performance to a classifier using age and BMI alone (Table 1).

Both the healthy vs pre-OA model and the healthy vs OA model were 94% accurate with an AUC of 0.96 and 0.99, respectively. Sensitivity and specificity were 1.0 and 0.90, respectively, for both the healthy vs pre-OA and healthy vs OA models. Classification of pre-OA using age and BMI alone was 83% accurate with an AUC of 0.78 and a sensitivity and specificity of 82% and 85%, respectively. Classification of OA using age and BMI alone was 59% accurate with sensitivity, specificity, and AUC of 67%, 55%, and 0.72, respectively. The AUC for both our pre-OA and OA models were higher than their corresponding classification models using age and BMI alone (pre-OA: p-value = 0.0476, CI = [0.004–0.402], OA: p-value = 0.004, CI = [0.097–0.476]) (Fig. 6D, Table 1).

The optimal weighting of scripted maneuvers for model performance was different between healthy vs pre-OA and healthy vs OA models. Healthy vs pre-OA was optimized with 40%, 35%, and 25% respective weights for flexion-extension, sit-to-stand, and walking. Healthy vs OA, however, was optimized predominantly with flexion-extension (70%) and, to a lesser extent, sit-to-stand (30%), with negligible contribution from walking measurements.

The models combining knee health status classes were also moderately successful (Supplementary Table 6 and 7). Our model classifying healthy knees from all grades of OA was 84% accurate with an AUC of 0.89. While the sensitivity of this model was high (96%), this configuration resulted in lower specificity (70%) due to struggling to classify healthy knees from older participants (of the 7 knees misclassified, 6 were healthy and classified as OA, average age of these participants was 66; the other was pre-OA, age 39, and classified as healthy). Our multi-class classification model for healthy, pre-OA, and OA knees was also less accurate than individual binary models (71% multi-class accuracy), with the model struggling most at classifying pre-OA (individual class accuracies of 75%, 55% and 83% for healthy, pre-OA, and OA respectively). Inclusion of available patient-reported outcomes, KOOS subscales as well as the presence or absence of crepitus, as features in our multiclass model was also investigated but did not improve classification accuracy (70% overall; 75%, 55% and 75% for healthy, pre-OA, and OA respectively).

E. Test-Retest Reliability of Knee Acoustics

Acoustic features were highly repeatable for all maneuvers and all 4 microphone locations. Intraclass and interclass correlation coefficients (ICC) for each feature and averaged between microphones were nearly all greater than 0.75 (Supplementary Table 8). ICC values were greater than 0.95 for all features selected for model optimization with recursive feature elimination. Arithmetic and geometric means were the only features with ICCs below our 0.75 repeatability threshold.

IV. Discussion

Our pilot study demonstrates the potential to use knee acoustics for identification and characterization of pre-OA in standard clinical settings. We measured KAE in acoustically

uncontrolled real-world settings using a portable device and used them to train and optimize a classification model which accurately classified knees with early OA as “injured” relative to “healthy” knees in our small cohort. This innovates on prior studies investigating KAE in osteoarthritis of the knee by demonstrating that KAEs can successfully differentiate osteoarthritic knees from healthy knees at an earlier stage of disease than has yet been reported and using a portable device at the point of care rather than in a controlled research setting.

It has been established that radiographic knee OA can be clearly differentiated from a healthy knee based on KAE [21], [22]. Our study replicated these observations. Other groups have had mixed success identifying radiographic knee OA in a study population of individuals with knee pain compatible with knee OA [20], [38]. Like these observations, which even with a larger sample size than ours had difficulty using acoustics to differentiate pre-radiographic from radiographic knee OA, our multi-class classification models for healthy vs all OA and healthy vs pre-OA vs OA generally produced poorer results than those of individual binary models. Clinically, however, the capacity to differentiate between stages of knee OA is not meaningful. Identification of individuals with early knee OA before it has progressed to the point that it is apparent on x-rays and without an expensive, time-consuming MRI is practical and an unmet clinical need. Thus, this was the primary strength and goal of our study. Identification of knee OA at an earlier stage of disease is a priority for OA research and clinical care because it may be more reversible or amenable to disease-modification with early and effective interventions.

Another strength of our study was inclusion of acoustics data from multiple different types of movement. The differential contribution of scripted maneuvers to and features selected for our optimized models suggest that the discriminatory information contained in KAE changes with the type and severity of underlying knee pathology in the context of knee biomechanics. Indeed, comparison of the performance of our classification models trained with the combined dataset vs those trained with individual maneuvers suggest that the combination of acoustic information from both non-weight-bearing and weight-bearing activities enhances model predictions. However, recording clean KAEs during dynamic, vigorous movements is non-trivial. The biomechanical complexity of weight-bearing movements like sit-to-stand and walking introduces the potential for noise. Variability between participant biomechanics due to physical fitness, kinematic adaptation to disease, or natural musculoskeletal variation could significantly alter acoustic features. Some of this variability can be accommodated for by selection or exclusion of movement cycles meeting certain criteria (i.e., pace or frequency) during signal processing or by utilizing measurements from specific anatomically placed microphones during model training. Future studies of KAE should include quantitative gait and movement measurements tightly coupled to KAE.

Additionally, the advantages of our design cannot be overstated. The portable nature of our device enables KAE measurements during multiple functionally relevant movements and at the point of care, which in the future would offer time savings for patients and immediate information for clinicians. Current modalities for evaluation of knee pain require a separate trip to a radiology suite, and in the case of MRI, a second thirty-to-sixty-minute

appointment typically on a different day and possibly at another location. The success of our classification algorithm in noisy environments also emphasizes the viability of KAE measurement in standard clinical environments. Further, the affordable nature of the design – approximately \$725 – grants far more accessibility to joint health evaluation compared to traditional imaging methods which require large up front capital expenditure as well as significant operating costs.

Future work will improve upon wearable device limitations such as system size and form factor. The piezoelectric contact microphones we used were effective at isolating KAEs in our noisy clinical environment and have been validated for joint health monitoring in previous studies [15], [27]. However, recent advancements in microphone technology and form factor could potentially improve microphone robustness during complex movements and improve KAE signal-to-noise ratio [39]. Additionally, the inclusion of IMUs in future designs would also eliminate the need for an external goniometer to estimate joint angle, improving system donning time and allowing for analysis in even more challenging environments [40]. Device form factor was not altered throughout the study for consistency, but our group also explored incorporating device hardware into an orthotic sleeve for intuitive donning of the system and improvements to cable management. Future work should investigate an orthotic design which can improve system usability without compromising signal quality with orthotic-derived noise artifacts.

Clinical limitations of our study include a small sample size, lack of a standardized definition of early knee OA, and cross-sectional design. It is highly likely that age, BMI, and contralateral knee OA, which are knee OA risk factors and uncontrolled for in our small cohort, influenced knee acoustics. Schluter et al. described both BMI and contralateral knee pain as influencing knee acoustics, albeit by different metrics than those we used [41]. While both our binary pre-OA and OA models using acoustic features significantly outperformed their counterpart models using age and BMI alone, age and BMI classification performance was still high for pre-OA relative to OA. Our pilot study size limited our ability to control for these covariates and thus we cannot rule out the possibility that our classification algorithm is implicitly capturing these variables to optimize performance. Further, we only collected KOOS on a subset of participants with pathologic knees. In a small pilot study, it would not be possible to robustly integrate patient-reported outcomes, such as KOOS, into our models because there is considerable heterogeneity in symptoms of knee OA cross-sectionally by radiographic stage as well as longitudinally in early knee OA [42], [43]. Future studies should be sufficiently powered to adjust for age, BMI, contralateral knee health status, and symptoms of knee OA, as well as to enable more data-intensive learning approaches. Next, there is currently not a universally accepted definition of early knee OA. Our pre-OA participants all had symptoms of knee OA (clinical classification criteria) and it is likely that these individuals, who did not have advanced enough radiographic changes to meet classification criteria for radiographic knee OA, are on the progressive spectrum of the natural history of knee OA and may not represent the earliest stages of disease or that is more reversible than established radiographic OA. Finally, while we feel our study is provocative in its identification of early disease, its cross-sectional nature precludes generation of any novel prognostic information. Identification of patients with pre-OA and a high likelihood of symptomatic or radiographic progression or transition to total knee

replacement is an unmet need in OA research. Further study of baseline or dynamic changes in acoustics on prediction of knee OA progression are warranted.

V. Conclusion

In summary, the results of our pilot study suggest that knee acoustic emissions measured with a wearable device at the point of care can accurately classify knees with pre-radiographic knee OA from healthy knees. Development of this technology could lead to improved access to tools for evaluation of knee health, lower costs, and potentially to better outcomes via earlier interventions for early disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

O. T. Inan is a co-founder of Arthroba Inc., a start-up focused on commercializing wearable joint health monitoring technologies. No other authors have competing interests.

This work was supported by VA Rehabilitation Research and Development (RR&D) as well as SPiRE II121RX003457-01A1, Grant #12040903

Biographies

Christopher J. Nichols studied at the University of Michigan and University of Cincinnati before completing his M.S. in biomedical engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA in 2022. Where he is currently pursuing his Ph.D. in bioengineering.

He joined the Inan Research Laboratory as a Graduate Research Assistant in 2020, where his research focuses on developing wearable technology and signal processing methods to expand musculoskeletal health monitoring technologies into everyday environments.



Goktug C. Ozmen received the B.S. and M.S. (Hons.) degrees in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2016 and 2019, respectively. He received the Ph.D. degree in electrical and computer engineering in 2022 from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA.

His research interests include design of wearable vibration and acoustics sensing systems for biomedical applications that enable non-invasive, reliable, and affordable physiological monitoring in uncontrolled settings. During his M.S. studies, he worked on design, simulation, and characterization of a MEMS membrane for a fiber optic microphone. Towards his Ph.D. degree with the Inan Research Laboratory, Georgia Tech, where he

has worked on projects involving the design of multi-modal wearable sensing systems for musculoskeletal health monitoring and signal processing methods using acoustic bio-signals for joint health assessment. His research interests include design of a wearable active vibration sensing system to enable non-invasive and robust assessment of knee health in field-deployable settings.

He received the Thesis of the Year Award from the Graduate School of Natural and Applied Sciences, METU, in 2019, with M.S. thesis entitled Design and Characterization of a MEMS Membrane for a Fiber Optic Microphone.



Kristine L. Richardson (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering, the M.S. degree in electrical and computer engineering, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2018, 2020, and 2022, respectively. Throughout her undergraduate degree, she interned with OFS Fitel LLC researching optical fiber data transmission and the Georgia Tech Research Institute researching in the field of biomedical engineering.

In 2019, she joined the Inan Research Laboratory as a Graduate Research Assistant. Her research concentrated on signal processing and machine learning for biomedical signals to assess joint health. She received the Szlam Fellowship in 2019 and the President's Fellowship from Georgia Tech in 2020. She now works as a researcher with Sandia National Laboratories.



Omer T. Inan (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 2004, 2005, and 2009, respectively. He joined ALZA Corporation (A Johnson and Johnson Company) in 2006, where he designed micropower circuits for iontophoretic drug delivery. In 2007, he joined Countryman Associates, Inc., Menlo Park, CA, USA, where he was the Chief Engineer, involved in designing and developing high-end professional audio circuits and systems. From 2009 to 2013, he was also a Visiting Scholar with the Department of Electrical Engineering, Stanford University. From 2013 to 2018, he was an Assistant Professor of Electrical and Computer Engineering with the Georgia Institute of Technology, where he is currently the Linda J. and Mark C. Smith Chair in Bioscience and Bioengineering, an Associate Professor of Electrical and Computer Engineering, and an Adjunct Associate Professor of Biomedical Engineering. He has published more than

230 technical articles in peer-reviewed international journals and papers in conferences and has eight issued patents. His research focuses on non-invasive physiologic sensing and modulation for human health and performance, including for chronic disease management, acute musculoskeletal injuries and disorders, and pediatric care.

Dr. Inan is an Invited Member of the IEEE Technical Committee on Translational Engineering for Healthcare Innovation and the IEEE Technical Committee on Cardiopulmonary Systems, and a technical program committee member or track chair for several other major international biomedical engineering conferences. He received the Gerald J. Lieberman Fellowship in 2009, the Lockheed Dean's Excellence in Teaching Award in 2016, the Sigma Xi Young Faculty Award in 2017, the IEEE Sensors Early Career Award in 2018, the Office of Naval Research Young Investigator Award in 2018, and the National Science Foundation CAREER Award in 2018. He also received an Academy Award for Technical Achievement from the Academy of Motion Picture Arts and Sciences (Oscars) in 2021. He was a National Collegiate Athletic Association (NCAA) All-American in the discus throw for three consecutive years (2001–2003). He is an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, the Theme 10 Editor of IEEE Engineering in Medicine and Biology Conference, and an Associate Editor of the IEEE Biomedical and Health Informatics Conference.



Dave Ewart was born in Chicago, IL, USA in 1984. He received his B.S.B. in finance from the University of Minnesota Carlson School of Management in 2008 and his MD from the University of Minnesota Medical School in 2012. He completed his internal medicine residency and rheumatology fellowship at the University of Minnesota and has been practicing as a rheumatologist at the Minneapolis Veterans Affairs Medical Center since 2018.



References

- [1]. Altman R et al. , "Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association," *Arthritis Rheum*, vol. 29, no. 8, pp. 1039–1049, Aug. 1986, doi: 10.1002/art.1780290816. [PubMed: 3741515]

- [2]. Ding C, Jones G, Wluka AE, and Cicuttini F, “What can we learn about osteoarthritis by studying a healthy person against a person with early onset of disease?,” *Curr. Opin. Rheumatol*, vol. 22, no. 5, pp. 520–527, Sep. 2010, doi: 10.1097/BOR.0b013e32833b90e9. [PubMed: 20592606]
- [3]. Karsdal MA et al. , “Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future,” *Osteoarthritis Cartilage*, vol. 24, no. 12, pp. 2013–2021, Dec. 2016, doi: 10.1016/j.joca.2016.07.017. [PubMed: 27492463]
- [4]. Katz JN et al. , “Disease modification in osteoarthritis; pathways to drug approval,” *Osteoarthr. Cartil. Open*, vol. 2, no. 2, p. 100059, Jun. 2020, doi: 10.1016/j.ocarto.2020.100059.
- [5]. Mahmoudian A, Lohmander LS, Mobasheri A et al., “Early-stage symptomatic osteoarthritis of the knee — time for action.”
- [6]. Hayashi D, Roemer FW, and Guermazi A, “Imaging of osteoarthritis-recent research developments and future perspective,” *Br. J. Radiol*, vol. 91, no. 1085, p. 20170349, May 2018, doi: 10.1259/bjr.20170349.
- [7]. Guermazi A et al. , “Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: population based observational study (Framingham Osteoarthritis Study),” *BMJ*, vol. 345, p. e5339, Aug. 2012, doi: 10.1136/bmj.e5339.
- [8]. Bacon K, LaValley MP, Jafarzadeh SR, and Felson D, “Does cartilage loss cause pain in osteoarthritis and if so, how much?,” *Ann. Rheum. Dis*, vol. 79, no. 8, pp. 1105–1110, Aug. 2020, doi: 10.1136/annrheumdis-2020-217363. [PubMed: 32381567]
- [9]. Horga LM et al. , “Prevalence of abnormal findings in 230 knees of asymptomatic adults using 3.0 T MRI,” *Skeletal Radiol*, vol. 49, no. 7, pp. 1099–1107, Jul. 2020, doi: 10.1007/s00256-020-03394-z. [PubMed: 32060622]
- [10]. Skou ST, Koes BW, Grønne DT, Young J, and Roos EM, “Comparison of three sets of clinical classification criteria for knee osteoarthritis: a cross-sectional study of 13,459 patients treated in primary care,” *Osteoarthritis Cartilage*, vol. 28, no. 2, pp. 167–172, Feb. 2020, doi: 10.1016/j.joca.2019.09.003. [PubMed: 31539621]
- [11]. Schiphof D, Waarsing EJ, Oei EH, and Bierma-Zeinstra SM, “Crepitus, joint line tenderness and the feeling of giving way are predictive signs for early knee osteoarthritis,” *Osteoarthritis Cartilage*, vol. 23, p. A330, Apr. 2015, doi: 10.1016/j.joca.2015.02.600.
- [12]. Lo GH, Strayhorn MT, Driban JB, Price LL, Eaton CB, and McAlindon TE, “Subjective Crepitus as a Risk Factor for Incident Symptomatic Knee Osteoarthritis: Data From the Osteoarthritis Initiative,” *Arthritis Care Res*, vol. 70, no. 1, pp. 53–60, 2018, doi: 10.1002/acr.23246.
- [13]. Mahmoudian A, Lohmander LS, Hawker G, Neogi T, and Englund M, “Association Of Baseline Knee Crepitus With Tibiofemoral Osteoarthritis Development In Patients With Knee Symptoms But No Or Minimal Radiographic Osteoarthritis Of The Knee,” *Osteoarthritis Cartilage*, vol. 31, pp. S214–S215, Mar. 2023, doi: 10.1016/j.joca.2023.01.194.
- [14]. Cibere J et al. , “Reliability of the knee examination in osteoarthritis: Effect of standardization,” *Arthritis Rheum*, vol. 50, no. 2, pp. 458–468, 2004, doi: 10.1002/art.20025. [PubMed: 14872488]
- [15]. Teague CN et al. , “A Wearable, Multimodal Sensing System to Monitor Knee Joint Health,” *IEEE Sens. J*, vol. 20, no. 18, pp. 10323–10334, Sep. 2020, doi: 10.1109/JSEN.2020.2994552.
- [16]. Töreyn H, Hersek S, Teague CN, and Inan OT, “A Proof-of-Concept System to Analyze Joint Sounds in Real Time for Knee Health Assessment in Uncontrolled Settings,” *IEEE Sens. J*, vol. 16, no. 9, pp. 2892–2893, May 2016, doi: 10.1109/JSEN.2016.2522964.
- [17]. Töreyn H, Jeong HK, Hersek S, Teague CN, and Inan OT, “Quantifying the Consistency of Wearable Knee Acoustical Emission Measurements During Complex Motions,” *IEEE J. Biomed. Health Inform*, vol. 20, no. 5, pp. 1265–1272, Sep. 2016, doi: 10.1109/JBHI.2016.2579610. [PubMed: 27305689]
- [18]. Gharehbaghi S, Jeong HK, Safaei M, and Inan OT, “A Feasibility Study on Tribological Origins of Knee Acoustic Emissions,” *IEEE Trans. Biomed. Eng*, vol. 69, no. 5, pp. 1685–1695, May 2022, doi: 10.1109/TBME.2021.3127030. [PubMed: 34757899]
- [19]. Scherpereel KL, Bolus NB, Jeong HK, Inan OT, and Young AJ, “Estimating Knee Joint Load Using Acoustic Emissions During Ambulation,” *Ann. Biomed. Eng*, vol. 49, no. 3, pp. 1000–1011, Mar. 2021, doi: 10.1007/s10439-020-02641-7. [PubMed: 33037511]

- [20]. Nevalainen MT et al. , “Acoustic emissions and kinematic instability of the osteoarthritic knee joint: comparison with radiographic findings,” *Sci. Rep.*, vol. 11, no. 1, Art. no. 1, Oct. 2021, doi: 10.1038/s41598-021-98945-2. [PubMed: 33414495]
- [21]. Prior J et al. , “Analysis of high frequency acoustic emission signals as a new approach for assessing knee osteoarthritis,” *Ann. Rheum. Dis.*, vol. 69, no. 5, pp. 929–930, May 2010, doi: 10.1136/ard.2009.112599. [PubMed: 20413570]
- [22]. Mascaro B, Prior J, Shark L-K, Selfe J, Cole P, and Goodacre J, “Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints,” *Med. Eng. Phys.*, vol. 31, no. 8, pp. 1013–1022, Oct. 2009, doi: 10.1016/j.medengphy.2009.06.007. [PubMed: 19643653]
- [23]. Kohn MD, Sassoon AA, and Fernando ND, “Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis,” *Clin. Orthop.*, vol. 474, no. 8, pp. 1886–1893, Aug. 2016, doi: 10.1007/s11999-016-4732-4. [PubMed: 26872913]
- [24]. Luyten FP et al. , “Toward classification criteria for early osteoarthritis of the knee,” *Semin. Arthritis Rheum.*, vol. 47, no. 4, pp. 457–463, Feb. 2018, doi: 10.1016/j.semarthrit.2017.08.006. [PubMed: 28917712]
- [25]. Roos EM and Lohmander LS, “The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis,” *Health Qual. Life Outcomes*, vol. 1, no. 1, p. 64, Nov. 2003, doi: 10.1186/1477-7525-1-64. [PubMed: 14613558]
- [26]. Jeong H-K, Pouyan MB, Whittingslow DC, Ganti V, and Inan OT, “Quantifying the Effects of Increasing Mechanical Stress on Knee Acoustical Emissions Using Unsupervised Graph Mining,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 3, pp. 594–601, Mar. 2018, doi: 10.1109/TNSRE.2018.2800702. [PubMed: 29522403]
- [27]. Richardson KL et al. , “Quantifying Rheumatoid Arthritis Disease Activity using a Multimodal Sensing Knee Brace,” *IEEE Trans. Biomed. Eng.*, pp. 1–1, 2022, doi: 10.1109/TBME.2022.3177074.
- [28]. Ozmen GC, Nevius BN, Nichols CJ, Mabrouk S, Teague CN, and Inan OT, “An Integrated Multimodal Knee Brace Enabling Mid-Activity Tracking for Joint Health Assessment,” in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Nov. 2021, pp. 7364–7368. doi: 10.1109/EMBC46164.2021.9630526.
- [29]. Nevius BN, “The Mechanical Design and Optimization of a Wearable Multimodal Health Sensing System,” Thesis, Georgia Institute of Technology, 2020. Accessed: May 25, 2022. [Online]. Available: <https://smartech.gatech.edu/handle/1853/64186>
- [30]. Ozmen GC, Safaei M, Lan L, and Inan OT, “A Novel Accelerometer Mounting Method for Sensing Performance Improvement in Acoustic Measurements From the Knee,” *J. Vib. Acoust.*, vol. 143, no. 3, Oct. 2020, doi: 10.1115/1.4048554.
- [31]. Khokhlova L, Komaris D-S, Tedesco S, and O’Flynn B, “Motion Artifact Resistant Mounting of Acoustic Emission Sensors for Knee Joint Monitoring,” in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Nov. 2021, pp. 7300–7303. doi: 10.1109/EMBC46164.2021.9629954.
- [32]. Whittingslow Daniel. C., Jeong H-K, Ganti VG, Kirkpatrick NJ, Kogler GF, and Inan OT, “Acoustic Emissions as a Non-invasive Biomarker of the Structural Health of the Knee,” *Ann. Biomed. Eng.*, vol. 48, no. 1, pp. 225–235, Jan. 2020, doi: 10.1007/s10439-019-02333-x. [PubMed: 31350620]
- [33]. Whittingslow DC et al. , “Knee Acoustic Emissions as a Digital Biomarker of Disease Status in Juvenile Idiopathic Arthritis,” *Front. Digit. Health*, vol. 2, 2020, Accessed: May 27, 2022. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fgth.2020.571839>
- [34]. Hall P and Wilson SR, “Two Guidelines for Bootstrap Hypothesis Testing,” *Biometrics*, vol. 47, no. 2, pp. 757–762, 1991, doi: 10.2307/2532163.
- [35]. Demler OV, Pencina MJ, and D’Agostino RB, “Misuse of DeLong test to compare AUCs for nested models,” *Stat. Med.*, vol. 31, no. 23, pp. 2577–2587, Oct. 2012, doi: 10.1002/sim.5328. [PubMed: 22415937]

- [36]. Koo TK and Li MY, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research," *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012. [PubMed: 27330520]
- [37]. Dwivedi D, Ganguly A, and Haragopal VV, "6 - Contrast between simple and complex classification algorithms," in *Statistical Modeling in Machine Learning*, Goswami T and Sinha GR, Eds., Academic Press, 2023, pp. 93–110. doi: 10.1016/B978-0-323-91776-6.00016-6.
- [38]. Kiselev J, Ziegler B, Schwalbe HJ, Franke RP, and Wolf U, "Detection of osteoarthritis using acoustic emission analysis," *Med. Eng. Phys.*, vol. 65, pp. 57–60, Mar. 2019, doi: 10.1016/j.medengphy.2019.01.002. [PubMed: 30685113]
- [39]. Lee SH et al. , "Fully portable continuous real-time auscultation with a soft wearable stethoscope designed for automated disease diagnosis," *Sci. Adv.*, vol. 8, no. 21, p. eabo5867, May 2022, doi: 10.1126/sciadv.abo5867.
- [40]. McGrath T and Stirling L, "Body-Worn IMU-Based Human Hip and Knee Kinematics Estimation during Treadmill Walking," *Sensors*, vol. 22, no. 7, Art. no. 7, Jan. 2022, doi: 10.3390/s22072544.
- [41]. Schlüter DK et al. , "Use of acoustic emission to identify novel candidate biomarkers for knee osteoarthritis (OA)," *PLOS ONE*, vol. 14, no. 10, p. e0223711, Oct. 2019, doi: 10.1371/journal.pone.0223711.
- [42]. Paradowski PT, Englund M, Roos EM, and Stefan Lohmander L, "Similar group mean scores, but large individual variations, in patient-relevant outcomes over 2 years in meniscectomized subjects with and without radiographic knee osteoarthritis," *Health Qual. Life Outcomes*, vol. 2, no. 1, p. 38, Jul. 2004, doi: 10.1186/1477-7525-2-38. [PubMed: 15279676]
- [43]. Roemer FW et al. , "From Early Radiographic Knee Osteoarthritis to Joint Arthroplasty: Determinants of Structural Progression and Symptoms," *Arthritis Care Res.*, vol. 70, no. 12, pp. 1778–1786, 2018, doi: 10.1002/acr.23545.

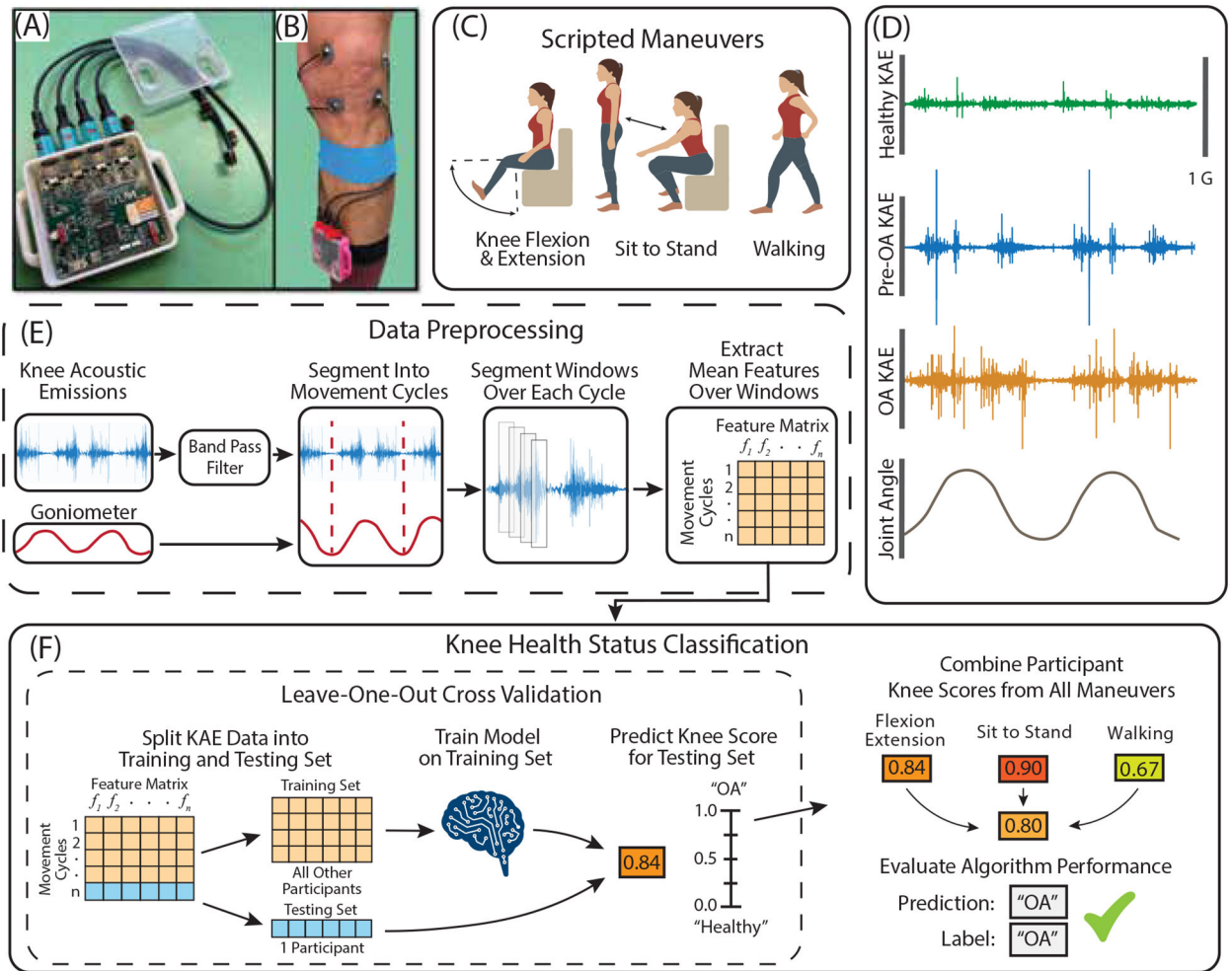
**Fig. 1:**

Illustration of hardware and acoustic data processing pipeline used in this study. (A) A custom printed circuit board designed by our group was integrated into a wearable form factor to record knee acoustic emissions. (B) The device was secured to the shin and microphones were attached to the knee with double sided adhesive and cables were secured with kinesiology tape. (C) Acoustic emissions were recorded from participants' knees while they performed three scripted maneuvers. (D) Example acoustic emissions data during flexion extension from healthy, pre-OA, and OA knees with similar demographics (Male, age 37, BMI 24.0 – 25.0). (E) Acoustic data were filtered and segmented into cycles either using a dual-axis goniometer or locating peaks in the acoustic signal envelope. Cycles were further subdivided into 30 millisecond windows with 20 millisecond overlap. Spectral, temporal, Mel-cepstral and band power features were then extracted from each window and averaged across each cycle to be used in evaluating classification models (E). Logistic regression with leave-one-out cross-validation was used to derive a 'Knee Score' denoting the likelihood a participant's knee was arthritic. Knee Scores were calculated for each scripted maneuver, and a composite Knee Score was calculated both by averaging scores across maneuvers as well as regressively weighting each

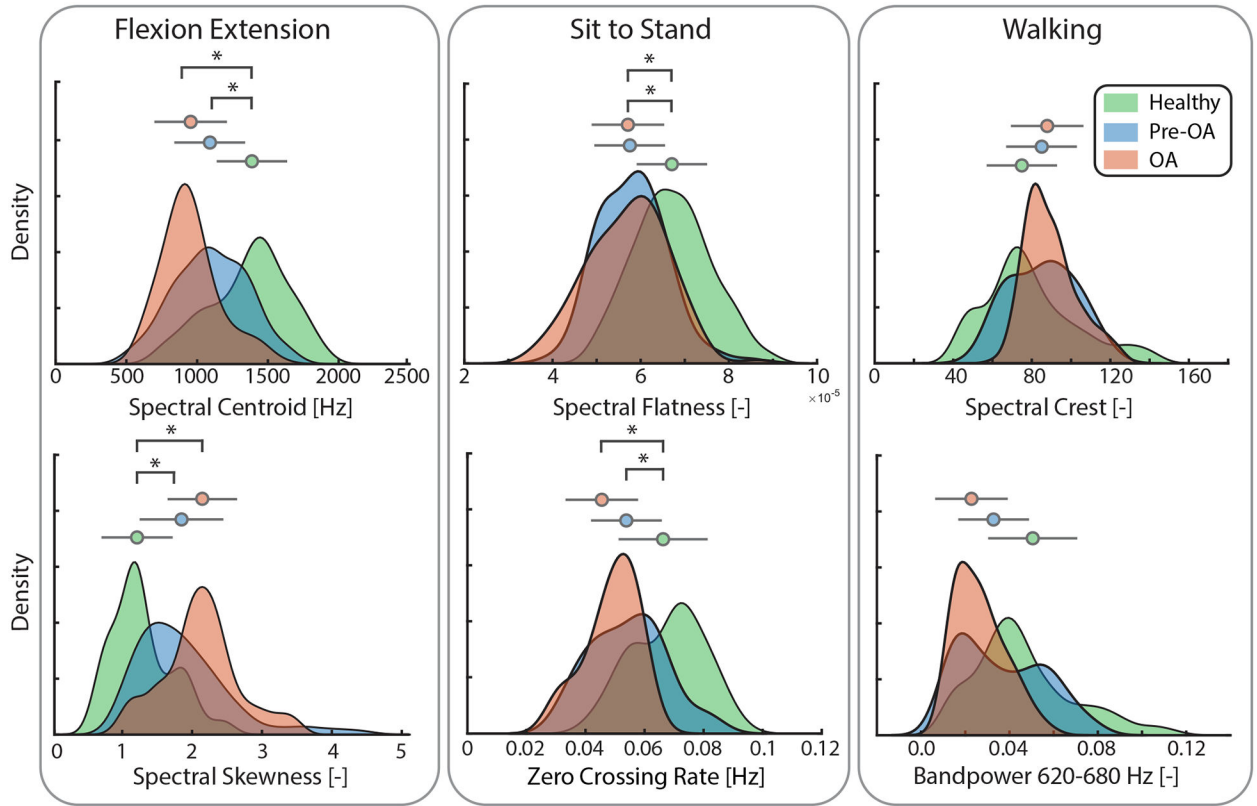
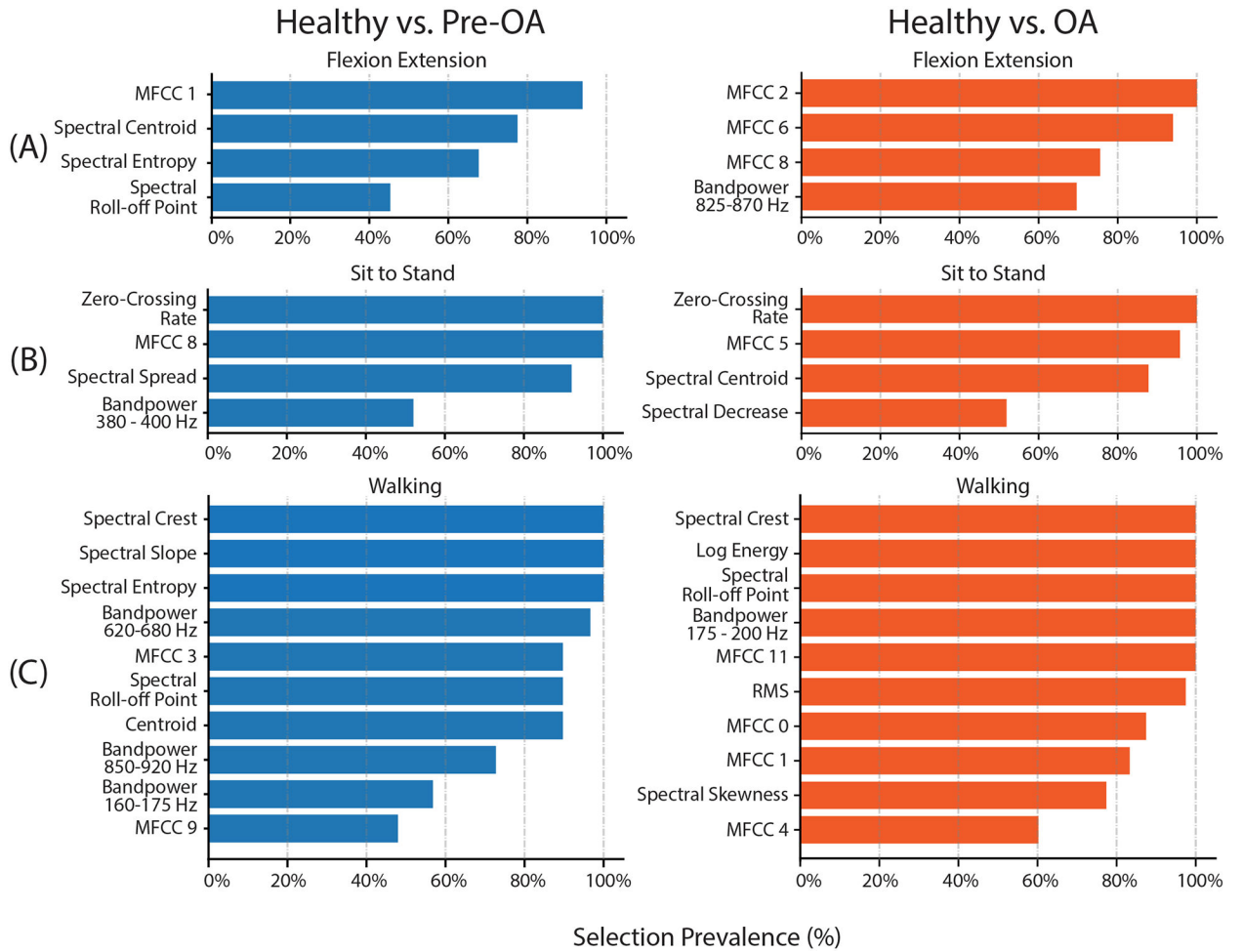


Fig. 2:

Groupwise density plots (distribution curves) and mean \pm standard deviation (circle and bars, respectively) for individual acoustic features selected for optimization of classification models and as measured during flexion-extension. Y axis denotes the probability density estimate using a gaussian kernel displaying the probability of feature distribution for each group. Acoustic features (x-axis) selected for classification model optimization during recursive feature elimination. Colored circles are the mean feature value, and the adjacent bars are the standard deviation. Feature-level visualization demonstrates how acoustic features change with increasing arthritic severity. [-]: unitless; [Hz]: hertz; *: p-value less than 0.025 with Dunn's post-hoc analyses; pre-OA: early pre-radiographic osteoarthritis; OA: radiographic osteoarthritis.

**Fig. 3:**

Feature selection prevalence from recursive feature selection during leave-one-out cross-validation of classification models. The number of features selected for each maneuver was optimized by increasing the number of included features until model accuracy improvement was saturated to avoid overfitting. Similar numbers of features were selected as optimal from each maneuver for healthy vs pre-OA and healthy vs OA models: 4 features each for flexion-extension, 4 features each for sit-to-stand, and 10 and 11 features each for walking. Features which were selected most frequently during cross-validation are highlighted. Feature selection across both models broadly emphasized consistent, low-frequency crepitus which was present in arthritic but not healthy knees. MFCC: Mel-Frequency Cepstrum Coefficient; RMS: root-mean-square value.

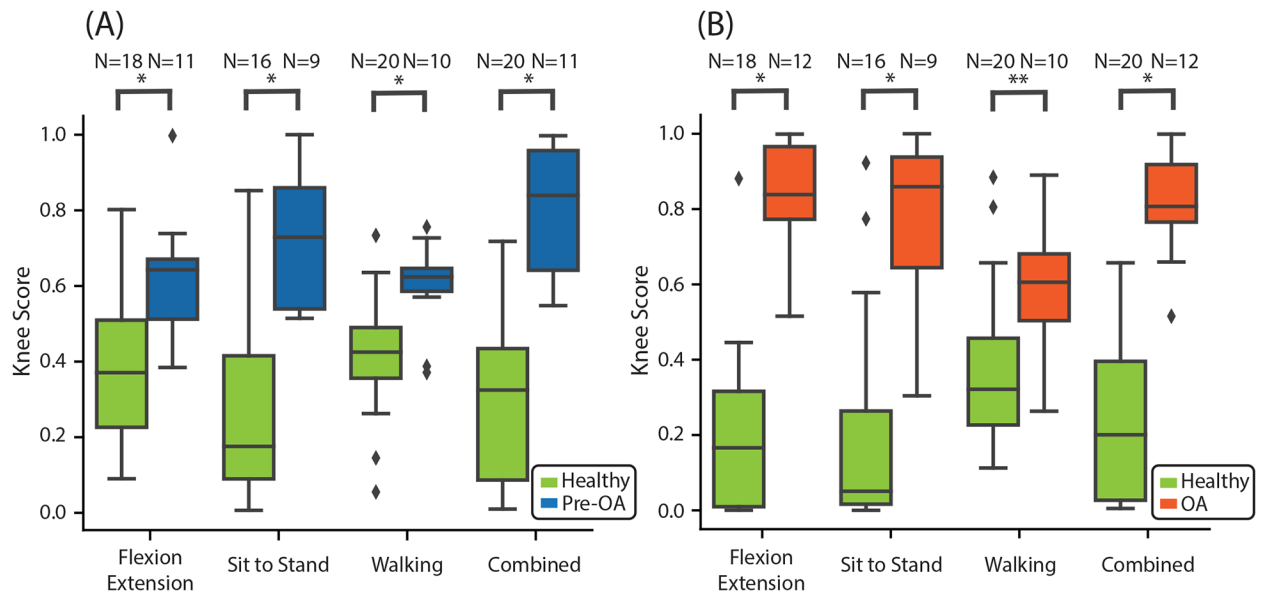


Fig. 4:

Model performance characteristics. Box-and-whisker plots of Knee Scores from our classifier comparing healthy and pre-OA (A) and healthy and OA (B). Knee Scores were lower across all maneuvers for healthy controls compared to pre-OA participants (FE: 0.38 ± 0.19 vs. 0.61 ± 0.16 , EF = 1.18; STS: 0.29 ± 0.27 vs. 0.74 ± 0.17 , EF = 1.73; walking: 0.41 ± 0.15 vs. 0.59 ± 0.12 , EF = 1.16; composite 0.37 ± 0.12 vs. 0.66 ± 0.13 , EF = 2.46; p-value < 0.005 for all comparisons), and for healthy controls compared to OA participants (FE: 0.20 ± 0.22 vs. 0.83 ± 0.15 , EF = 3.01; STS: 0.20 ± 0.29 vs. 0.78 ± 0.22 , EF = 1.98; walking: 0.38 ± 0.20 vs. 0.60 ± 0.18 , EF = 1.07; composite: 0.240 ± 0.22 vs. 0.82 ± 0.13 , EF = 3.17; p-value < 0.005 for all maneuvers except walking, where the p-value was 0.007). *: p-value less than 0.005, **: p-value = 0.007, EF: effect size (Hedge's *g*), diamond: outlier, pre-OA: early pre-radiographic osteoarthritis; OA: radiographic osteoarthritis; FE: flexion-extension, STS: sit-to-stand.

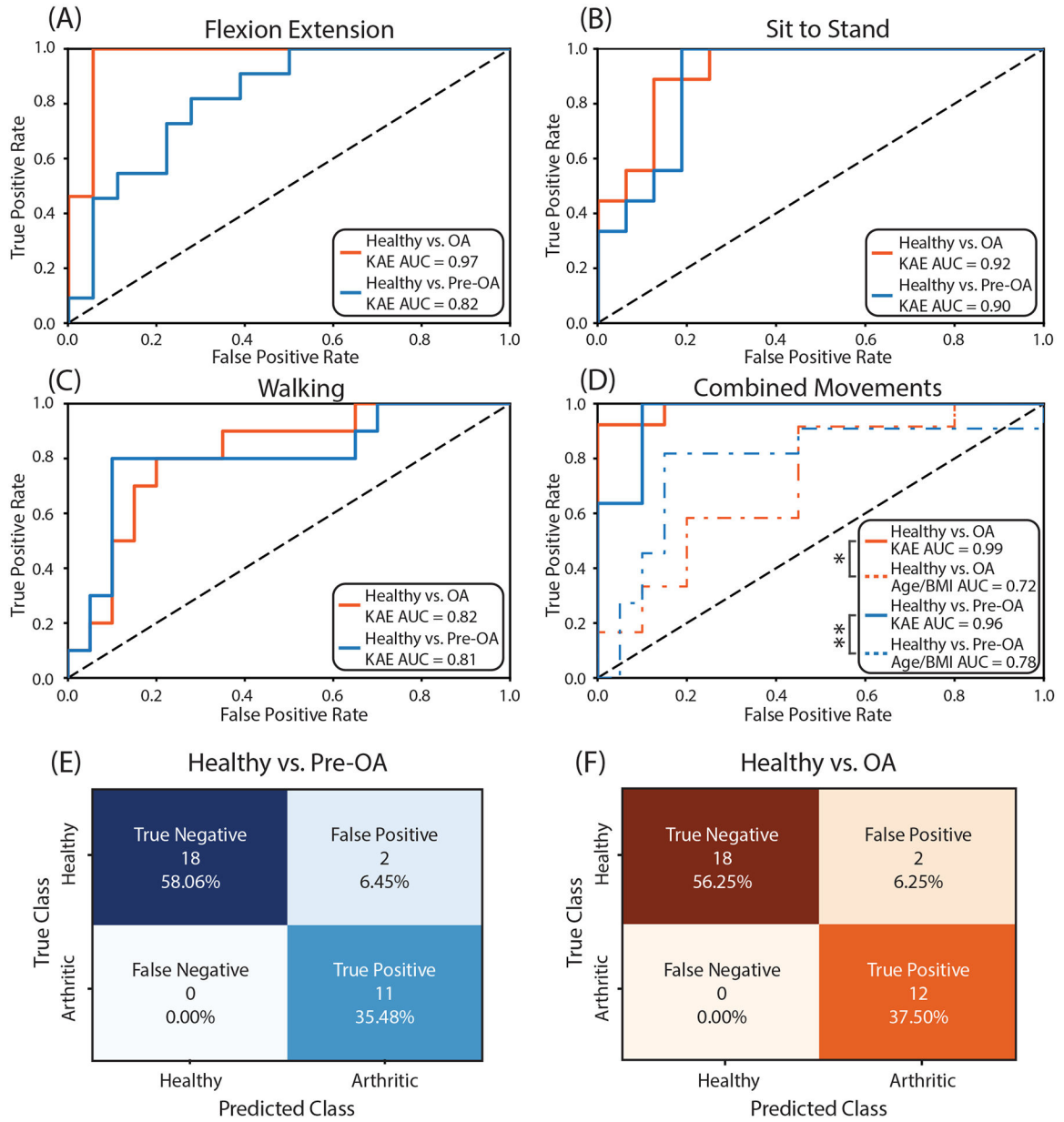


Fig. 5: Receiver operating characteristic curves for flexion-extension (A), sit-to-stand (B), walking (C), and the composite of all three maneuvers compared to classification using age and BMI (D) for healthy vs pre-OA and healthy vs OA. Sit-to-stand had the highest AUC of the three individual scripted maneuvers for our healthy vs pre-OA model (AUC = 0.88) while flexion-extension elicited the best individual performance for our healthy vs. OA model (AUC = 0.97). Models trained with combined maneuver dataset had higher absolute AUC values than models trained with individual maneuvers (healthy vs. pre-OA AUC = 0.96, healthy vs. OA AUC = 0.99). Composite models using KAEs also had higher AUC values than models using age and BMI for classification (healthy vs. pre-OA AUC = 0.78, healthy vs. OA AUC = 0.72). (E) Confusion matrix for composite healthy and pre-OA

classification. (F) Confusion matrix for composite healthy and OA classification. pre-OA: early pre-radiographic osteoarthritis; OA: radiographic osteoarthritis; AUC: area under the curve; FE: flexion-extension, STS: sit-to-stand; KAE: Knee acoustic emissions.

TABLE I

Healthy vs. Pre-OA Classification Performance

Movement Type	Healthy Mean \pm S.D.	Arthritic Mean \pm S.D.	P value	Effect Size (Hedge's g)	CI 95%	ROC AUC	Accuracy	Sensitivity	Specificity
Flexion-Extension	0.38 \pm 0.19	0.61 \pm 0.16	<0.005	1.18	[0.09 0.37]	0.82	76%	82%	72%
Sit-to-Stand	0.29 \pm 0.27	0.74 \pm 0.17	<0.005	1.73	[0.25 0.64]	0.90	88%	100%	81%
Walking	0.41 \pm 0.15	0.59 \pm 0.12	<0.005	1.16	[0.07 0.29]	0.81	80%	80%	80%
Composite	0.30 \pm 0.20	0.81 \pm 0.15	<0.005	2.46	[0.37 0.65]	0.96	94%	100%	90%
Age / BMI	0.36 \pm 0.20	0.65 \pm 0.25	0.006	1.26	[0.05 0.43]	0.78	83%	82%	85%